I) Review and preview

II) Loans, bonds, interest rates and yields

III) Rate of return and interest-rate risk

IV) Interest rate differentials

V) Term structure

VI) Liquidity

VII) Loans

VIII) Money, liquidity preference and determination of interest rates in a simple economy

IX) Financial intermediation

X) Central banks: introduction

XI) The central bank as lender of last resort (LOLR)

XII) Financial crises and Fed actions in recent years

XIII) How a central bank controls the overnight interest rate

XIV) Monetary policy and macroeconomics, simplified

XV) Monetary policy and macroeconomics, complications

©**Christopher Hanes   Economics 450   Class notes**

II) Loans, bonds, interest rates and yields

A) Intro

1) What we will cover in this section

Here I will tell you about two ways to borrow and lend money: loans and bonds. Probably, you already know something about loans but you don't know about bonds. I will do the usual thing in an economics class: I will start

with the simplest possible case and gradually add complications. I will make things only as complicated as needed for you to understand monetary policy, so there are many aspects of bonds and loans that I won't get to.

I will be introducing a lot of *notation*, that is a system of symbols used to describe things. Pay attention to the notation, even though it is boring. We will be using it all through the course.

2) Some things to remember from middle-school math

a) Percents

A percentage is a fraction multipled by 100, so

"Ten percent" or "10%" means a decimal or fraction of $0.10 = \frac{1}{10}$

"Five percent" or "5%" means $0.05 = \frac{1}{20}$

"Half a percent" or "1/2%" means $0.005 = \frac{1}{200}$

b) Powers and exponents

You need to remember how to use powers and exponents. I mean this stuff:

$$x^{1/2} = \sqrt{x} \qquad x^{1/3} = \sqrt[3]{x} \qquad \sqrt{x^2} = x \qquad \left(x^2\right)^{1/2} = x \qquad (x^2)^{1/3} = x^{2/3} \qquad x^{-1} = \frac{1}{x}$$

$$x^{-1/2} = \frac{1}{x^{1/2}} = \frac{1}{\sqrt{x}}$$

If you don't remember this stuff, go to a middle school and ask a seventh-grader about it. The best middle school around here is Vestal Middle School, at 600 South Benita Boulevard, Vestal NY.

B) Definition: borrowers and lenders

I use the word "borrower" to mean a person or firm, or government, or whatever that wants to get money today, and in return will give a promise to pay out a specified amount of money at a specified time in the future.

A "lender" is the opposite: a lender is willing to give money today to acquire a promise of a specified amount of money at a specified future point in time.

Transactions between borrowers and lenders can be structured in many different ways, but there are two main ways: loans, and bonds.

What about shares of stock? A sale of stock is not a borrowing/lending transaction, on our definition. When you are buying a share of stock, you are buying a share of a firm's profits. The directors of the firm choose whether to pay those profits to you ("dividends") or reinvest them in the firm ("retained earnings"). In the latter case, you might eventually benefit from getting more dividends in the future or by selling your shares to someone else, but you get no payments in the meantime. Thus, when you buy a share of stock you do *not* receive "a promise of a specified amount of money at a specified future point in time." We won't talk about stocks much in this course.

C) One-year loans and bonds

1) Introduction

The simplest case to describe is one where the borrower's promise is to make *one* payment of money exactly *one year from today*. So I'll start with that. I should mention, however, that in reality many borrowers actually *do* borrow money for exactly one year. Retailers borrow to buy stocks of goods, promising to pay money back in a year from the proceeds of sales. Manufacturers borrow to buy materials and components, promising to pay money back in a year from the proceeds of selling what they make. Governments borrow money for a year to fund expenditures, promising to pay money back when annual tax payments come in.

2) One-year loans

In a one-year loan, the lender gives the borrower some money today and the borrower gives the lender a promise that the borrower will pay the lender a certain amount of money one year from today. Usually, the amount of money that borrower promises to pay the lender is specified in terms of an *interest rate*, rather than a number of dollars.

Here's an example: you borrow $100 from your father today, for one year, at an interest rate of 5 percent (5%). That means that today your father gives you $100. You give him a promise that one year from today you will pay him $100 plus five percent of $100. Five percent of $100 is $5, so you will pay your father $105.

Often, the terms of a loan are put into a written contract. Such a contract would name the lender, the borrower, the amount the borrower receives today, and the interest rate.

Here's the notation we'll use to describe one-year loans.

L is the amount of money the borrower gives the lender today.

i is the interest rate expressed as a decimal or fraction. I hope you paid attention to that. If I say "the interest rate on the loan is one percent," then i in the formula is 0.01 *not* 1.

$(1+i)\$L = \$L + \$iL$ is the amount of money that the borrower promises to give the lender one year from today.

In the example of the loan from your father, L is 100 and i is 0.05.

Sometimes L is called the "principal" of the loan, and iL is called the "interest." The total $(1+i)\$L$ is "principal plus interest."

An important ratio to think about, for what we will do later, is: <u>The future payment promised by the borrower</u>
The amount of money the lender pays out today

Using our notation, this ratio is $\dfrac{(1+i)L}{L} = 1+i$

3) One-year bonds

a) What they are

Like a loan contract, a bond is a legally-binding piece of paper. But it is not a contract between a lender and a borrower. It is a sort of one-way contract binding on a borrower. It is a piece of paper on which a borrower promises to pay a certain amount (or amounts) of money, at a specified point (or several points) in time, *to whomever is holding the piece of paper at that point in time.* It names the borrower, the date on which the borrower promises to pay money to the holder of the bond, and the amount of money that the borrower promises to pay on that date.

3

The borrower prints up the piece of paper, makes it legally binding - in the old days, this was by signing the piece of paper - and *sells* the piece of paper for as much as he can get. This is called "issuing" the bond. The borrower is called the "issuer" of the bond.

I said a bond is a "piece of paper." That isn't quite right nowadays. All bonds used to be printed (or written) on paper. Nowadays most bonds are just records in a database - virtual pieces of paper. But I don't like electronic things, so I'll go on talking about piece of paper.

There are many ways that a borrower/issuer can sell a bond, but for right now, suppose that the issuer runs an auction, or puts the bond up on Ebay and takes the highest bid. The issuer knows exactly how much money he is promising to pay in the future - he has printed that on the bond. But he does *not* necessarily know what price he will get for the bond when he sells it. That depends on the result of the auction, or the offers he gets on Ebay.

There is no price and no name of a lender written on a bond. The borrower/issuer does not know who will buy the bond, or what price the buyer will pay, until the auction/Ebay sale is complete.

Here's the notation we'll use to describe a one-year bond.

IOU is the amount of money the bond issuer promises to pay to the bondholder one year from today.

P is the price someone pays for the bond today.

For an example, suppose the General Motors Corporation has issued a bond that promises to pay $150,000 one year from today, and somone just bought that bond for $135,000. Then *IOU = 150,000* and *P = 135,000*.

Importantly, the person who buys the bond from the issuer is not necessarily the person to whom the issuer will pay the *IOU* when the bond comes due. The person who buys the bond from the issuer can sell the bond to *another* person. The second person might sell the bond on to a third person, and so on. A bond can change hands many times before it comes due. Whoever is holding the bond when it comes due will get the IOU payment from the borrower/issuer.

In this way, a bond is different from a loan. If a lender makes a one-year loan to someone, the lender must wait a year, the time specified on the loan contract, to get any money back. (I am simplifying things a bit here; I'll return to this later.) If a lender buys a one-year bond, on the other hand, the lender has more options. She can wait a year to get the *IOU, or* she can sell the bond right now to someone else. If she sells the bond right now, she does not get the *IOU*: she gets whatever someone else is willing to pay for the bond at the time she sells it. That price might be more, or less, than the price she paid for the bond earlier. I'll return to this point later.

b) Yield: the implicit interest rate on a bond

I haven't mentioned an interest rate. In fact, there is no interest rate written on a bond. There is, however, a sort of *implicit* interest rate defined by the *IOU* and *P*. This implicit interest rate is called a "yield." We will denote it with an *i*, the way we denote interest rates.

It is defined like this:

$$i = \frac{IOU}{P} - 1 \quad \text{or} \quad 1 + i = \frac{IOU}{P}$$

In the example of the General Motors bond,

$$i = \frac{150,000}{135,000} - 1 = 1.1111 - 1 = 0.1111$$

Like interest rates on loans, bond yields are usually spoken of in terms of percents, not fractions. So one would say that the yield on this bond is "eleven percent" (doing some rounding).

I called the yield an "implicit interest rate." Indeed, the yield on a bond is called the "interest rate" on the bond. Why? How is a yield like an interest rate?

Remember the ratio I mentioned in reference to loans, that is: $\dfrac{\text{The future payment promised by the borrower}}{\text{The amount of money the lender pays out today}}$

For a loan this ratio was $\dfrac{(1+i)L}{L} = 1+i$

For a bond this ratio is $\dfrac{IOU}{P} = 1+i$

Did you see how the yield on a bond is like the interest rate on a loan? Like the interest rate on a loan, a bond's yield expresses the ratio of the future payment made by the borrower to the amount a lender invests today.

If I tell you a bond's IOU and its yield, you can figure out what its price is:

$$P = \frac{IOU}{1+i}$$

This is important! In bond markets, people usually don't quote prices of bonds. They quote bond yields, from which infer the prices. A financial website giving information about bonds will quote yields not prices.

Finally, note that if I tell you a bond's price and yield, you can figure out what its IOU is:

$$IOU = P(1+i)$$

c) Changes in bond prices and yields

Unlike the interest rate on a loan, the yield on a bond can change from day to day. The yield on a bond is defined by the last price that someone paid for a bond. If the market price of the bond changes - if someone pays a different price for the bond today than someone paid for the bond yesterday - the bond's yield changes.

By definition, if a bond's market price rises, its yield falls, and vice-versa. And remember that bond yields are often called interest rates. So these three headlines which you might see on a financial-news website all mean the same thing:

"Bond prices fall'
"Bond yields rise"
"Interest rates rise"

d) Market prices adjust to equalize bond yields

To a first approximation, market prices of bonds must be such that yields are the same for all bonds that pay off at the same point in time. This isn't *exactly* true, as I will explain later, but for now, for simplicity, assume it is true.

Why would this be true? It is an equilibrium condition. The profit-maximizing actions of traders in bond markets ensures that it holds.

Suppose there are two one-year bonds on the market, bond A and bond B. Bond A's yield is determined by its market price and *IOU*:

$$i_A = \frac{IOU_A}{P_A} - 1$$

Bond B's yield is: $i_B = \frac{IOU_B}{P_B} - 1$

Suppose that bond B's price is so high, relative to its IOU, that bond B's yield is *lower* than bond A's: that is $i_B < i_A$. This situation cannot prevail for long! It means that bond B is a *bad deal* for potential bond buyers: bond A gives more return per dollar invested. So people rush to sell bond B and buy bond A. As people sell bond B, its market price falls, which raises its yield, making it a better deal than it was before:

$$i_B \uparrow = \frac{IOU_B}{P_B \downarrow} - 1$$

Meanwhile, as people rush to buy bond A its market price rises, reducing its yield and making it a worse deal than it was before;

$$i_A \downarrow = \frac{IOU_A}{P_A \uparrow} - 1$$

In equilibrium, the two bonds must be equally good deals, so their yields must be the same. The actions of investors in bond markets ensure that the bonds' market prices adjust until their yields are equal.

How long does it take to get to this equilibrium? No more than a few minutes. Bond traders watch market bond prices very carefully. Nowadays bond traders run computer programs that place buy and sell orders whenever bond prices are out of whack.

4) Using interest rates and yields: some examples

a) Inferring a bond's price from its yield

When people talk about a stock's price, they usually talk about the price per share. But when people talk about a bond's price, they do it by talking about the yield. If you went to a financial website to find out about one-year bonds being bought and sold today, you'd see quotes of yields, not prices.

From the yield and the amount of the bond's *IOU*, you can calculate the price. Consider a one-year bond with an *IOU* of $5,000. If you read that the yield on this bond is 2.5 percent, you know that the market price of the bond is:

$$P = \frac{IOU}{1+i} = \frac{5,000}{1+0.025} = \frac{5,000}{1.025} = 4,878.05$$

b) What price is an issuer likely to get for a newly issued bond?

Suppose a firm wants to borrow money by issuing one-year bonds. It prints up a mess of bonds, each with an *IOU* of $10,000, planning to auction them off. What price is it likely to receive for each bond?

To figure that out, the firm can look on a financial website and see what the yields have been on one-year bonds sold this morning. Because yields must be the same for all one-year bonds, the price the firm will get when it auctions off its newly-issued one-year bonds will be the price that equates the yield on those bonds to the yields on other one-year bonds being sold this morning.

Suppose yields on one-year bonds sold this morning have been 1/2%. Then the price the firm will receive for each of its bonds is likely to be:

$$P = \frac{IOU}{1+i} = \frac{10,000}{1+0.005} = \frac{10,000}{1.005} = 9,950.25$$

c) Rolling over a series of one-year loans

Suppose that you make a one-year loan. The principal on the loan - the amount of money you give the borrower today - is L. When the borrower pays the money back a year from now, you take all of what she pays you - principal plus interest - and make *another* one-year loan (maybe to the same borrower, maybe to a different borrower). Two years from now, you are paid the principal plus interest on that second one-year loan. Again you take all that money and make *yet another* one-year loan. And so on until, after several years, you cash out.

Making a series of short-term loans like this, always lending out all the principal and interest from the previous loan, is called "rolling over" short-term loans.

How much money will you have when you cash out?

To describe this, we must take allow for the possibility that the interest rate you can get on a one-year loans may not be the same next year, or the year after that, as it is today. In fact, it is not just a possibility, it is a probability - interest rates on newly negotiated one-year loans rarely remain exactly the same from one day to the next, even.

Here's the notation we will use:

i_t is today's interest rate on one year loans.

i_{t+1} is the interest rate on one year loans that will prevail one year from today.

i_{t+2} is the interest rate on one-year loans that will prevail two years from today.

and so on.

Thus, if you start by making a loan of L,
at the end of the first year, you will have $(1+i_t)L$.

at the end of the second year, you will have $(1+i_{t+1})(1+i_t)L$

(Remember, you took *all* the money you got back at the end of the first year, which was $(1+i_t)L$, and lent all of that out again.)

at the end of the third year, you will have $(1+i_{t+2})(1+i_{t+1})(1+i_t)L$

and so on.

Finally, let's imagine a case which is unrealistic - it would almost never happen in reality - but which will prove useful for a point I will be making later on. Assume the one-year interest rate remains exactly the same year after year. Let's denote this unchanging interest rate by $\hat{i}$. So:

$i_t = \hat{i}$

$i_{t+1} = \hat{i}$

$i_{t+2} = \hat{i}$

and so on. Thus,

at the end of the first year, you will have $(1+\hat{i})L$.

at the end of the second year, you will have $(1+\hat{i})(1+\hat{i})L = (1+\hat{i})^2 L$

at the end of the third year, you will have $(1+\hat{i})(1+\hat{i})(1+\hat{i})L = (1+\hat{i})^3 L$

and so on.

Notice that in this special case, because the interest rate is the same in all years, we can write your final cash-out amount in a different way. If we denote the number of years before you cash out by n, we can say that the amount of money you will have when you cash out is $(1+\hat{i})^n L$.

d) Rolling over a series of one-year bonds

Let's do the rollover story again but with bonds this time.

You start with $\$X$ today. You spend all of the $\$X$ on one-year bonds. Perhaps you buy just one bond with a price that happens to equal $\$X$ exactly. Perhaps you buy two or more bonds whose prices total to $\$X$.

At the end of the year you are paid the *IOU*'s on those bonds. You take all of that money, the total of all the IOU's, and buy another mess of one-year bonds. When those bonds pay off, on a date exactly two years from now, you take the total of the *IOU*'s on *those* bonds and buy *another* mess of one-year bonds.... And so on until, after several years, you cash out. Buying a series of short-term bonds like this, reinvesting all of the *IOU*'s each time, is called "rolling over" short-term bonds.

How much money will you have when you cash out?

To figure this out, recall that, by definition, the *IOU* on a one-year bond is equal to the bond's price times (one plus the yield): $IOU = (1+i)P$, where i stands for the yield on one-year bonds. Thus, whatever specific combination of bonds you buy, as long as you spend $\$X$ in total the total of the IOU's you will be paid at the end of the year is $(1+i)X$. And $(1+i)X$ is the total amount of money you will spend buying your second round of one-year bonds a year from now.

But wait a minute! Like prevailing interest rates on one-year loans, the market yield on one-year bonds can change from day to day. It almost never remains exactly the same from one year to the next. To allow for this, use the same notation I used in the previous section:

i_t is today's yield on one-year bonds.

i_{t+1} is the yield on one-year bonds that will prevail one year from today.

i_{t+2} is the yield on one-year bonds that will prevail two years from today.

and so on.

Thus,

at the end of the first year, you will be paid IOU's that total $(1+i_t)X$.

at the end of the second year, you will be paid IOU's that total $(1+i_{t+1})(1+i_t)X$

at the end of the third year, you will be paid IOU's that total $(1+i_{t+2})(1+i_{t+1})(1+i_t)X$

and so on.

Finally, we can again imagine an unrealistic case in which the yield on one-year bonds does not change, it remains equal to $\hat{i}$ year after year. In that case,

at the end of the first year, you will have $(1+\hat{i})X$.

At the end of the second year, you will have $(1+\hat{i})(1+\hat{i})X = (1+\hat{i})^2 X$

At the end of the third year, you will have $(1+\hat{i})(1+\hat{i})(1+\hat{i})X = (1+\hat{i})^3 X$

and so on.

Let n denote the number of years before you cash out. Then we can generally say that the amount of money you will have when you cash out is $(1+\hat{i})^n X$.

D) More complicated bonds and loans

1) Introduction

Now let's make things more complicated. Bonds and loans can be more complicated in a number of ways. The date on which the borrower promises to make the payment might not be exactly one year from today: it might be two or more years in the future (or less than a year). The borrower might promise to make multiple payments, on multiple dates - that is one payment one year from now, another payment two years from now, and so on. In a case where the borrower promises multiple payments, the first payment might be the same amount of money as the second payment, or more money, or less money, and so on.

I will call bonds and loans in which the borrower promises to make just one payment "single-payment" loans and bonds. Loans and bonds in which the borrower promises to make a series of payments are "multiple-payment."

2) Maturity

The amount of time between now and the specified date on which the borrower will pay off the interest and principal on a loan, or pay the *IOU* on a bond, is called the "maturity" of the loan or bond.

Thus, the bonds and loans I described in the previous section all had a maturity of exactly one year.

A single-payment loan or bond in which the borrower promises to pay the interest and principal on the loan, or pay the IOU on the bond, *two* years from today has a maturity of two years. And so on.

For multiple-payment bonds and loans, "maturity" means the amount of time between now and the date specified for the *final* payment.

Here is something important: the maturity of a bond *changes over time* over the lifetime of the bond. By "lifetime of the bond" I mean the years between its issuance and when the borrower pays the last IOU.

For example, suppose that in September of 2020 the General Motors Corporation issues (prints up and sells) a bond that will make one payment in September of 2030. On the day it is issued, the maturity of that bond is ten years, because its IOU will pay off ten years in the future. As of September 2021, however, one would say that the maturity of the bond is nine years. As of September 2022, the maturity of the bond is eight years. In September 2029, it will be a one-year bond.

The maturity of a bond at the time it is issued is called its ***original maturity***.

(Something you don't need to know, but I'll tell you anyway: sometimes, but not often, people reserve the term "maturity" to mean original maturity and use the term "tenor" to refer to the amount of time between now and when a bond or loan pays off. Anyway, if you see the word "tenor" used in reference to financial markets, that's what it means.)

3) Single-payment loans of maturity greater than one year

In this type of loan the lender gives the borrower some money today and the borrower gives the lender a promise that the borrower will pay the lender a specified amount of money two years, or three years, or four years....from today, as the case may be. As in a one-year loan, the amount of money that borrower promises to pay the lender is specified in terms of an interest rate rather than a number of dollars.

It is customary to define all loans' interest rates on an "annual" or "annualized" basis, even if the maturity of the loan is more or less than one year. What does this mean? It is easiest to understand with examples.

Consider a loan in which the lender gives the borrower $L today. The loan contract says that the borrower will make one payment to the lender two years from today and the interest rate - that means annualized interest rate - is five percent. What does this mean? It means that on the date two years from now the borrower will pay the lender an amount of money equal to:

$$(1+0.05)(1+0.05)L = (1+0.05)^2 L$$

Another example. A loan contract says that the borrower will pay three years from today and the interest rate is one percent. This means is that three years from now the borrower will pay the lender an amount of money equal to:

$$(1+0.01)(1+0.01)(1+0.01)L = (1+0.01)^3 L$$

To generalize, if the single loan payment will be made m years from now and the annualized interest rate is i, the amount of the payment will be:

$$(1+i)^m L$$

Notice that this looks like the formula we derived in part IIIC4c). If you think about, you can see that the "annualized interest rate" on a single-payment loan of maturity greater than one year is the answer to the question: what would the interest rate on rolled-over one-year loans have to be in order to give the lender the same final payoff as this longer-maturity loan? Read that again and again until you understand it.

And of course, for $m=1$, this is the same formula we used for one-year loans.

4) Single-payment bonds of maturity greater than one year

a) What they are

This type of bond is a piece of paper on which the issuer/borrower promises to pay an IOU to whoever is holding the piece of paper on a date two or more years from today.

b) Yield to maturity

As in the case of a one-year bond, one speaks of the "yield" on the bond, which is defined by the ratio of the IOU on the bond relative to the price paid for the bond in the latest sale of the bond. Just as it is conventional to define all loans' interest rates on an annualized basis, it is conventional to define all bonds' yields on an annualized basis. A yield defined in this way is called, specifically, the "yield to maturity." There are other kinds of yields, but you can usually assume that "yield" means "yield to maturity."

What is this "yield to maturity"? I need to introduce some more notation here.

$_m i$ denotes the yield to maturity on a single-payment bond with a maturity of m years. That is to say, it is a bond on which the issuer is supposed to pay the IOU exactly m years from now.

For example, the yield to maturity on a bond that pays off one year from now is denoted $_1i$. The yield to maturity on a bond that pays off four years from now is denoted $_4i$. The yield to maturity on a bond that pays off 117 years from now is $_{117}i$.

In this notation, here is the formula that defines yield to maturity on a single-payment bond:

$$_mi = \sqrt[m]{\frac{IOU}{P}} - 1 = \left(\frac{IOU}{P}\right)^{1/m} - 1$$

This formula gives the yield as a fraction or decimal, which must be multipled by 100 to give the yield as a percent. Note that, for $m=1$, this is the same as the definition I gave earlier for the yield on a one-year bond.

Here's an example. A bond will pay an IOU of $20,000 two years from now. Its market price today is $18,000. The yield (yield to maturity) on this bond is:

$$i = \sqrt{\frac{20,000}{18,000}} - 1 = \left(\frac{20,000}{18,000}\right)^{1/2} - 1 = 1.054 - 1 = 0.054 \text{ or } 5.4 \text{ percent.}$$

As with one-year bonds, when people talk about prices of longer-maturity bonds they do it by talking about the yield. If you went to a financial website to find out about longer-maturity bonds being bought and sold today, you'd see quotes of yields (that is, yield to maturity), not prices.

5) Multiple-payment bonds

a) General definition

This means a bond in which the issuer promises to make one payment one year from today, another payment two years from today, another payment three years from today, and so on.

In most *real* multiple-payment bonds, the payments come once a quarter. (A "quarter" means a quarter of a year, three months). For simplicity, we'll just pretend the payments come once a year.

Our notation:
$_1IOU$ is the first payment
$_2IOU$ is the second payment
$_3IOU$ is the third payment
and so on.

There are multiple-payment loans, too, but we don't deal with those much in this course, so I won't introduce notation for them.

Back when bonds were pieces of paper, a multiple-payment bond was several pieces of paper, one for each *IOU*, batched together. You can imagine them as stapled together. (Really, they were printed on one big sheet with perforations, like a sheet of postage stamps.) When the first *IOU* came due, the holder of the bond would detach it from the bunch and present it to the issuer to get paid. When the second *IOU* came due, the holder would detach that *IOU* and present it to the issuer, and so on until all the *IOU*'s were gone.

The issuer sells the *IOU*'s as a bunch, not separately, getting one price for the whole bunch.

But after the bond has been sold by the issuer, anyone holding a multiple-payment bond is free to *pull the IOU's apart and sell them individually* as single-payment bonds. This is important. I'll get back to it later.

b) Yield to maturity on a multiple-payment bond

When I was talking about single-payment bonds, I said people in bond markets usually don't quote prices for bonds directly; they quote yields and leave it to you to calculate what the price must be, and that "yield" means an annualized yield called "yield to maturity." All that is true for multiple-payment bonds as well.

What is the yield to maturity for a multiple-payment bond? Like the yield to maturity of a single-payment bond, it is a number defined by the market price of the bond, the maturity of the bond and its *IOUs*. You get it by setting up and solving an equation. But the equation for a multiple-payment bond is complicated. It is so complicated that you can't solve it algebraically! You have to let a computer or financial calculator solve it for you.

I'll use $_{YTM}i$ to denote the yield to maturity on a multiple-payment bond. It is a   For a bond that makes two payments, the equation is:

$$P = \frac{{}_1IOU}{(1 + {}_{YTM}i)} + \frac{{}_2IOU}{(1 + {}_{YTM}i)^2}$$

For a bond that makes four payments,

$$P = \frac{{}_1IOU}{(1 + {}_{YTM}i)} + \frac{{}_2IOU}{(1 + {}_{YTM}i)^2} + \frac{{}_3IOU}{(1 + {}_{YTM}i)^3} + \frac{{}_4IOU}{(1 + {}_{YTM}i)^4}$$

and so on.

To generalize, to calculate the yield to maturity on a multiple-payment bond of n years maturity, you take as given the *IOU's* on the bond and the market price of the bond. The yield to maturity is the value of $_{YTM}i$ that solves this equation:

$$P = \frac{{}_1IOU}{(1 + {}_{YTM}i)} + \frac{{}_2IOU}{(1 + {}_{YTM}i)^2} + \ldots\ldots + \frac{{}_nIOU}{(1 + {}_{YTM}i)^n}$$

As an example, consider a bond that makes payments for four years. For the first three years the IOU is $500. For the last year the IOU is $1000. The market price of the bond is $1879. This bonds yield to maturity is the value for $_{YTM}i$ that solves:

$$1879 = \frac{500}{(1 + {}_{YTM}i)} + \frac{500}{(1 + {}_{YTM}i)^2} + \frac{500}{(1 + {}_{YTM}i)^3} + \frac{1000}{(1 + {}_{YTM}i)^4}$$

As I said, this is not a math problem you can solve algebraically. You have to plug the numbers into a financial calculator or computer and let the machine iterate until it finds the solution. For this example, the solution is $_{YTM}i = 0.11$ or 11 percent

c) What is the most you would be willing to pay for a multiple-payment bond?

A multiple-payment bond is really just a bunch of single-payment bonds stuck together. You could get exactly the same thing by buying a bunch of single-payment bonds that give the same pattern of *IOU*'s over time. Consider the bond in the example above. You could reproduce this bond by buying a one-year single-payment bond with an *IOU* of $500, along with a two-year single-payment bond with an *IOU* of $500, a three-year bond with an *IOU* of $500, and a four-year bond with a $1000 *IOU* (or two four-year bonds with *IOU*'s of $500 each).

The fact that you can reproduce any multiple-payment bond by buying a set of single-payment bonds establishes a *maximum* that you would be willing to pay for any multiple-payment bond. That maximum is the total cost of buying the set of single-payment bonds that reproduces the multiple-payment bond. As an analogy, suppose that at the grocery store you patronize a jar of mayonnaise costs $2.50, a bottle of ketchup is $3, and a jug of milk is $4. What is the most you would be wiling to pay for a bag containing a jar of mayonnaise, a bottle of ketchup and a jug of milk? $9.50!

Let's apply this logic to the bond in the example above. How much would it cost to reproduce this bond with a batch of single-payment bonds? Remember that bond dealers quote prices for bonds indrectly, by quoting yields. You can look at a financial website that gives you current yields on single-payment bonds. Suppose the website says that current market yields on zero-coupon bonds are as follows:

$$_1i = 0.01 \ (one \ percent)$$

$$_2i = 0.04 \ (four \ percent)$$

$$_3i = 0.02 \ (two \ percent)$$

$$_4i = 0.03 \ (three \ percent)$$

Then:

a zero-coupon bond that will pay you $500 in one year costs $P = \dfrac{500}{(1+0.01)} = 495.05$

a zero-coupon bond that pays $500 in two years costs $P = \dfrac{500}{(1+0.04)^2} = 462.28$

a zero-coupon bond paying $500 in three years costs $P = \dfrac{500}{(1+0.02)^3} = 471.16$

and a bond paying $1000 in four years costs $P = \dfrac{1000}{(1+0.03)^4} = 888.49$

The total cost of *all* these bonds would be *495.05 + 462.28 + 471.16 + 888.49 = 2316.98*
So $2316.98 is the most you, or anyone, would be willing to pay for the multiple-payment bond in question.

To generalize, to calculate the most a buyer would be willing to pay for a multiple-payment bond with a maturity of *n* years, you take as given the bond's *IOU*'s and current yields on single-payment bonds, and use this formula to calculate the maximum price anyone would pay:

$$P = \frac{_1IOU}{(1+_1i)} + \frac{_2IOU}{(1+_2i)^2} + \ldots\ldots + \frac{_mIOU}{(1+_mi)^m}$$

WARNING! It is easy for you to confuse this formula with the one I gave above that defines the yield to maturity on a multiple-payment bond. They are not the same. Make sure that you understand how they are different.

d) What is the market price of a multiple-IOU bond?

The price we just talked about in the previous section, that is the highest price anyone would be willing to pay for a multiple-payment bond, is also the *equilibrium market price* of the bond. The profit-maximizing actions of traders and dealers in bond markets ensure that is the case.

To understand why, you must remember something I said earlier: "after the bond has been sold by the issuer, the person who bought the bond can pull all the IOU's apart and sell them individually as single-payment bonds." And you must consider what would happen if the market were out of equilibrium, that is if the market price were greater than/less than the equibrium price defined in the previous section.

What would happen if the market price were less than the equilibrium price, that is if:

$$P < \frac{{}_1 IOU}{(1+{}_1 i)} + \frac{{}_2 IOU}{(1+{}_2 i)^2} + \ldots\ldots + \frac{{}_m IOU}{(1+{}_2 i)^m}$$

In this case, you could profit money by buying up as many of these particular multiple payment bonds as you can get your hands on. You take each bond, pull all of its *IOUs* apart into separate pieces of paper which are now single-payment bonds. You sell each of these single-payment bonds individually. What is the total amount that you will receive when you sell them individually? Well the price of ${}_1 IOU$ will of course be determined by the current yield on one-year single-payment bonds, ${}_1 i$. The price of ${}_2 IOU$ will be determined by ${}_2 i$, and so on. The total amount you will receive is:

$$\frac{{}_1 IOU}{(1+{}_1 i)} + \frac{{}_2 IOU}{(1+{}_2 i)^2} + \ldots\ldots + \frac{{}_m IOU}{(1+{}_m i)^m}$$

As long as this number is greater than the multiple-payment bond's price P, you will make a profit. This process is called **stripping** a bond, because you are stripping apart the components of the multiple-payment bond to make many single-payment bonds.

Traders and dealers in bond markets are quick to take advantage of profitable opportunities. So you can be sure that they will in fact rush to buy these particular multiple-IOU bonds, to strip them. That will drive up the market price of the multiple-payment bond until the profit opportunity - the price disequilibrium - has disappeared. This makes money for you, because the total you will receive by selling the IOUs as single-payment bonds is equal to: the right-hand side of the expression above.

What if the opposite case holds? What would happen if the market price were more than the equilibrium price, like:

$$P > \frac{{}_1 IOU}{(1+{}_1 i)} + \frac{{}_2 IOU}{(1+{}_2 i)^2} + \ldots\ldots + \frac{{}_m IOU}{(1+{}_m i)^m}$$

In that case, traders and dealers in bond markets would do the *opposite of stripping*. They would buy a batch of single-payment bonds that reproduce the payment pattern of the overvalued multiple-IOU bond, thus creating more bonds identical to the overvalued bond, and sell those newly-created multiple-payment bonds at a profit. Their actions would, of course, drive *down* the market price of the multiple-payment bond until the price disequilibrium disappears. There is not a special word for this process, but it happens.

E) Names of common bond types

1) Introduction

I have used the word "bond" to refer to all instruments of borrowing and lending that feature *IOU*'s payable to whoever is holding the *IOU* at the time it comes due. I will continue to do that in this course, to keep our terms simple. But in reality, there are different names for different types of "bonds." In this section I will tell you some of those names. Keep in mind that many bonds that do not fall into these common types. Almost any pattern of *IOU*'s you can imagine has been issued as a bond at one time or another.

Finally, remember what the phrase "original maturity" means: the maturity of a bond at the time it is issued.

2) Types of single-payment bonds

a) Bill

A single-payment bond with an original maturity of one year or less is call a "bill."

A "Treasury bill" is a bill issued by the U.S. Treasury.

"Commercial paper" is a bill issued by a private corporation.

b) Zero-coupon bond

This is a single-payment bond with an original maturity greater than one year. You'll come to understand why it is called this when you read below about "coupon bonds."

c) Strip

This is a zero-coupon bond that was originally part of a multiple-payment bond, and was stripped.

2) Types of multiple-payment bonds

a) Fixed-payment bond

A fixed-payment bond is a bond that makes one payment every year for *m* years - that is, its maturity is *m* years - and each payment is the same size (same amount of money). Example: a bond that pays $100 once a year for fifty years.

b) Coupon bond

A coupon bond is a bond that makes a small payment once a year for *n* years, all the same size (same amount of money), and in the final year - the *m*th year - makes one *big* payment in addition to the small payment. An example is a bond that pays an *IOU* of $50 every year for five years, and in the fifth year pays a $50 *IOU* plus *another* IOU of $1,000.

A coupon bond's small payments are called "coupons" or "coupon payments." We will denote them by *C*. In the example above, *C = $50*.

The extra, big payment that comes in the last year is called the "par value" or "face value" of the bond. We will denote it by *F*. In the example above, *F = $1,000*.

The *coupon rate* is C/F. This is a fraction, usually spoken of as a percent. In the example, the coupon rate is:

$$\frac{C}{F} = \frac{50}{1000} = 0.05 \text{ or 5 percent.}$$

c) Consol (*or* perpetuity)

This is a fixed-payment bond that makes a payment once a year forever - the maturity is infinity. Such bonds were issued in the nineteenth century by the governments of Britain and the U.S. They are called "consols" after a British government issue of 1752. At that time the British government had many bonds outstanding, issued at various times in the past. It was a pain for the Treasury to manage the payment of the various *IOU*'s. So the Treasury issued a new set of bonds and used the money from the sale of these bonds to buy up all the outstanding bonds. An operation like this is called "consolidating" debt. The new bonds were perpetuities. So "consol," from "consolidated," became a word for any perpetual bond.

©**Christopher Hanes   Economics 450   Class notes**

III) Rate of return and interest-rate risk
      A) Introduction
      B) Rate of return: definition
            1) Single-payment bond
                  a) Held for one year
                  b) Held for *n* years
            2) Coupon bond held for one year
      C) Holding to maturity *versus* selling before maturity
      D) "Interest-rate risk" (or "duration risk")
            1) Definition
            2) Interest-rate risk for bonds of short *versus* long maturities
            3) How to avoid interest-rate risk
            4) Duration

III) Rate of return and interest-rate risk

A) Introduction

Recall that an interest rate measures the return that a lender receives on a loan. A yield measures the return that a lender receives on a bond held to maturity (so that the lender gets paid the IOU). But a lender can sell a bond before it matures. How do we describe the return a lender receives in that case?

By a number called the "rate of return." The rate of return depends on the price the lender receives when she sells the bond, relative to the price she paid for the bond. Like interest rates and yields, rates of return are conventionally expressed on an annualized basis.

Unlike interest rates and yields, rates of return are uncertain until the bond is sold. Market yields, hence bond prices, change from day to day. Some of those changes are predictable, but many are not. Thus, a lender planning to sell a bond before it matures does not know for sure what price she will receive when she sells the bond. At the time she buys the bond, she can *guess* what that price will be but she can't know for *sure*.

A rate of return can easily be a negative number. A rate of return is negative if the price at which the lender sells the bond is lower than the price at which she bought the bond. Of course, that is not what a lender intended to happen when she bought the bond. But it can turn out to happen if future yields turn out to be much higher than the lender expected them to be when she bought the bond.

The formula that defines the rate of return depends on how long the lender held the bond, and whether the bond is a coupon bond or a single-payment bond (that is a "bill" or "zero-coupon bond").

B) Rate of return: definition

1) Single-payment bond

a) Held for one year

I start by defining the rate of return if a lender buys a single-payment (zero-coupon) bond that pays off more than one year from today - that is, its maturity *m* is greater than one - but the lender sells the bond after just one year. Make sure you remember: this lender does *not* get paid the *IOU*.

I need to use the notation I introduced earlier to keep track of the passage of time, in which the subscript *t* denotes "today," *t+1* is "a year from now," *t+2* is "two years from now" and so on.

$_mP_t$ is the price today of a bond with a maturity of m years.

$_{m-1}P_{t+1}$ is the price of that same bond when it is sold one year from today, at time $t+1$. Note that at time $t+1$ (next year) the maturity of the is $m-1$, not m.

The rate of return received by someone who buys the bond today and sells it a year from today, denoted RER, is:

$$RER = \frac{_{m-1}P_{t+1}}{_mP_t} - 1 = \frac{_{m-1}P_{t+1}}{_mP_t} - \frac{_mP_t}{_mP_t} = \frac{_{m-1}P_{t+1} - _mP_t}{_mP_t}$$

This is a fraction, spoken of as a percent. For example, if I buy a bond for $100 and sell it next year for $300, the rate of return is:

$$RER = \frac{300}{100} - 1 = 2 \text{ or 200 percent.}$$

A rate of return can be negative. For example, if I buy a bond today for $100 and sell it next year for $50, the rate of return is:

$$RER = \frac{50}{100} - 1 = -\frac{1}{2} \text{ or -50 percent.}$$

b) Held for n years

What is the rate of return if a lender buys a single-payment (zero-coupon) bond and sells it after n years, but still before the bond matures? Again remember this lender will not get paid the IOU.

The price when the lender buys the bond is $_mP_t$. The price when the lender sells the bond is $_{m-n}P_{t+n}$.

$$RER = \sqrt[n]{\frac{_{m-n}P_{t+n}}{_mP_t}} - 1 = \left(\frac{_{m-n}P_{t+n}}{_mP_t}\right)^{1/n} - 1$$

Again this is a fraction spoken of as a percent. For example, if I buy a bond today for $169 and sell it three years from now for $201, the rate of return is:

$$RER = \sqrt[3]{\frac{201}{169}} - 1 = \left(\frac{201}{169}\right)^{1/3} - 1 = 1.06 - 1 = 0.06 \text{ or 6 percent.}$$

Anothere example: if I buy a bond today for $2 aand sell it two years from today for $8, the rate of return is:

$$RER = \sqrt{\frac{8}{2}} - 1 = (4)^{1/2} - 1 = 2 - 1 = 1 \text{ or 100 percent.}$$

2) Coupon bond held for one year

For a coupon bond things are more complicated. The seller of the bond won't get the face value (par value), but as long as she holds the bond for at least a year she will get at least one coupon payment. Recall that we denote the amount of a coupon payment by C.

To keep things simple, we will only consider a case in which the lender buys a coupon bond and sells it after one year. We won't consider cases in which the bond is sold after two or more years - the formula for that is a mess!

The rate of return to buying a coupon bond and selling it after one year is:

$$RER = \frac{_{m-1}P_{t+1} + C}{_mP_t} - 1 = \frac{_{m-1}P_{t+1} + C}{_mP_t} - \frac{_mP_t}{_mP_t} = \frac{_{m-1}P_{t+1} + C - _mP_t}{_mP_t}$$

(again a fraction, spoken of as a percent).

C) Holding to maturity *versus* selling before maturity

Here is an important thing to notice. Someone who is planning to sell a bond before it matures, or at least thinks that he *might* want to sell the bond before it matures, cares a lot about what happens to prices in bond markets. If he goes onto a financial-news website and reads a headline that says:

"Bond prices fell yesterday"

"Yields rose yesterday" (remember, those two headlines mean the same thing)

he is sad: it means he will get less for his bond when (or if) he sells it.

Headlines that say:

"Bond prices fell yesterday"

"Yields rose yesterday"

make him happy: it means he will get more for his bond when (or if) he sells it.

What about a person who is planning to hold a bond until it matures, and is *sure* he won't have to sell the bond before that? He does not care about those headlines. He does not care what happens to bond prices while he is holding the bond. He is going to get the *IOU*, no matter what happens to bond prices in the meantime.

D) "Interest-rate risk" (or "duration risk")

1) Definition

Interest-rate risk is one of the big ideas in this course. It will come up in several places.

It is a risk faced by a person who holds a bond and is planning to sell a bond before it matures, or might have to do so depending on circumstances.

It is the risk that an unpredicted increase in market yields will reduce the price at which the person could sell the bond. Of course, there is an upside too: an unpredicted decrease in market yields would increase the bond's sale price. But the situation is *risky* compared with making a loan or holding a bond to maturity. In both those cases, the lender knows exactly how much money he will be paid in the future.

Why is this risk called "interest-rate risk"? Remember that bond yields are often called "interest rates," because yields are implicit interest rates on bonds. Hence interest-rate risk. Perhaps a better name for it would be "unpredictable-future-yield risk," or at least "yield risk." But that is not what it is called.

A common alternative name for interest-rate risk is "duration risk." I'll explain this later on.

2) Interest-rate risk is greater for bonds of longer maturity
What does this statement mean?

Suppose that market yields go up unexpectedly for bonds of all maturities, by exactly the same amount. For an example, say that all yields increase by two percent: yields on one-year bonds $_1i$ go up by 2 percent; yields on two-year bonds $_2i$ also go up by 2 percent; yields on three-year bonds $_3i$ *also* go up by 2 percent; and so on.

Prices of bonds at all maturities will fall. (Indeed, to say that yields go up is the same as saying the yields fall.) You already know this.

But here's something you probably don't know: prices will fall *more*, percentagewise, for bonds of *longer* maturity. That is, the percent decrease in prices of two-year bonds will be bigger than the percent decrease in prices of one-year bonds; the percent decrease in prices of three-year bonds will be bigger than the percent decrease in prices of two-year bonds; and so on.

Yes, there is an upside too: if market yields fall unexpectedly for bonds of all maturities, by exactly the same amount at all maturities, then prices will rise more, percentagewise, for bonds of longer maturity. But the point is that longer-maturity bonds are *riskier* in this respect.

There are several ways to see that an increase in yields has a larger percentage effect on prices of longer-maturity bonds, if you fiddle around with the formula that defines the relationship between bond prices and yields. But in this class we'll just demonstrate the point by example.

In the example, we will consider two people. One of them buys a two-year single-payment (zero-coupon) bond and sells it after one year. The other one buys a three-year bond and sells it after one year. (Note that the three-year bond is the one of longer maturity, so that bond should be more subject to interest-rate risk.) In both cases, yields rise unexpectedly, so the rate of return on the transaction turns out to be lower than the person expected. For the example, I will make sure that the unexpected increase in yields is exactly the same for both bonds. What you will see is that the reduction in the rate of return is bigger - worse - for the three-year bond.

Here's how we describe the two cases in notation.

Buy a two-year bond and sell it after one year

Price a person pays for the bond today: $_2P_t = \dfrac{IOU}{(1 + {_2i_t})^2}$

Price the person receives when the bond is sold: $_1P_{t+1} = \dfrac{IOU}{(1 + {_1i_{t+1}})}$

Rate of return: $_2RER = \dfrac{_1P_{t+1}}{_2P_t} - 1 = \dfrac{_1P_{t+1} - _2P_t}{_2P_t}$

Buy a three-year bond and sell it after one year

4

Price a person pays for the bond today: $_3P_t = \dfrac{IOU}{(1+\,_3i_t)^3}$

Price the person receives when the bond is sold: $_2P_{t+1} = \dfrac{IOU}{(1+\,_2i_{t+1})^2}$

Rate of return: $_3RER = \dfrac{_2P_{t+1}}{_3P_t} - 1 = \dfrac{_2P_{t+1} - _1P_t}{_1P_t}$

Now, let's put in specific numbers. The *IOU* on both bonds is $1000. At the time the bonds are purchased, market yields for two-year bonds are 5 percent. Market yields for three-year bonds are also 5 percent.

What both bond buyers *expect* is that bond yields will remain 5 percent in the future. That's not what turns out to happen, but let's figure out what they were expecting to receive.

What they were expecting to receive

- Buy a two-year bond and sell it after one year

Price a person pays for the bond today: $_2P_t = \dfrac{IOU}{(1+\,_2i_t)^2} = \dfrac{1000}{(1+0.05)^2} = 907$

Price the person receives when the bond is sold: $_1P_{t+1} = \dfrac{1000}{(1+0.05)} = 952$

Rate of return: $_2RER = \dfrac{_1P_{t+1}}{_2P_t} - 1 = \dfrac{_1P_{t+1} - _2P_t}{_2P_t} = \dfrac{952 - 907}{907} = 0.05$ or 5%

- Buy a three-year bond and sell it after one year

Price a person pays for the bond today: $_3P_t = \dfrac{1000}{(1+0.05)^3} = 864$

Price the person receives when the bond is sold: $_2P_{t+1} = \dfrac{IOU}{(1+\,_2i_{t+1})^2} = \dfrac{1000}{(1+0.05)^2} = 907$ or 5%

Rate of return: $_3RER = \dfrac{_2P_{t+1}}{_3P_t} - 1 = \dfrac{_2P_{t+1} - _1P_t}{_1P_t} = \dfrac{907 - 864}{864} = 0.05$ or 5%

What actually happens? Yields for both bonds jump to 10 percent!

What they actually receive

- Buy a two-year bond and sell it after one year

Price a person pays for the bond today: $_2P_t = \dfrac{IOU}{(1+\,_2i_t)^2} = \dfrac{1000}{(1+0.05)^2} = 907$

Price the person receives when the bond is sold: $_1P_{t+1} = \dfrac{1000}{(1+0.10)} = 909$

Rate of return: $_2RER = \dfrac{_1P_{t+1}}{_2P_t} - 1 = \dfrac{_1P_{t+1} - _2P_t}{_2P_t} = \dfrac{909 - 907}{907} = 0.002$ \qquad or 0.2%

- Buy a three-year bond and sell it after one year

Price a person pays for the bond today: $_3P_t = \dfrac{1000}{(1+0.05)^3} = 864$

Price the person receives when the bond is sold: $_2P_{t+1} = \dfrac{IOU}{(1+\,_2i_{t+1})^2} = \dfrac{1000}{(1+0.10)^2} = 826$ or 5%

Rate of return: $_3RER = \dfrac{_2P_{t+1}}{_3P_t} - 1 = \dfrac{_2P_{t+1} - _1P_t}{_1P_t} = \dfrac{825 - 864}{864} = -0.04$ or - 4%

Look! The person who bought the longer maturity bond, the three-year bond, has been hurt worse by the inexpected increase in yields, even though the increase in yields was the same for both maturities!

Any example you set up in which yields increase by the same amount for bonds of different maturities will give the same result: there is a bigger loss in the return to holding the longer-maturity bond.

3) How to avoid interest-rate risk

Remember that a person who is certain that he will hold a bond until maturity doesn't care what happens to the price of the bond after he buys it. Thus, one can avoid interest-rate risk by never selling a bond before it matures.

But that is not so easy. To do it, you have to know exactly the future points in time at which you will need money, and buy bonds at the maturities that exactly match that timing.

For some lenders (investors), and some future expenditures, that is possible. Consider, for example, a man who is saving to pay for his young childrens' future college tuition. (I am describing myself years ago. I am now very, very old. My youngest child has finished college. But once, in the almost unimaginably distant past, I was saving for college tuition.) The amount of that expenditure is not perfectly predictable, but the timing of it is. If a person buys bonds for that purpose, it is pretty easy to match up the maturity of the bonds to the maturity of the expenditure.

But for many lenders and many future expenditures it is not possible. The timing of some expenditures is unpredictable. Consider, for example, the expenditure of buying a new car, for me at least. I drive a car until it is old and unreliable. When it breaks down and the repair would cost more than the car is worth, I junk the car and buy a new one. I can't predict exactly when this will happen.

These examples were silly, but you get the point: it is not always possible to predict when an expenditure will have to be made. Many lenders have to worry about interest-rate risk on longer-maturity bonds.

4) Duration

The word "duration," in the context of bonds markets, refers to a number that measures the degree of interest-rate risk on a bond. If the "duration" of bond A is greater than the "duration" of bond B, then the price of bond A is more strongly affected by a given change in yields.

I said earlier that "duration risk" was another term for interest-rate risk. Now you see why.

For a single-payment (zero-coupon) bond, duration is simple. It is the same as the maturity of the bond: the number of years from now until the time the bond pays its single IOU.

For multiple-payment bonds, duration is more complicated. A multiple-payment bond is has some IOUs that will be paid in the near future and some IOU's that will be paid in the further future - it is sort of a combination of short-term bonds and long-term bonds. Thus, for a multiple-payment bond the number we call "duration" is a sort of average of the maturities of the many IOU's making up the bond.

The average is complicated, however, so that is all you need to know any more about "duration" for this class.

©**Christopher Hanes   Economics 450   Class notes**

IV) Interest rate differentials
        A) Introduction
        B) Default risk
                1) Definition of default risk
                2) Default risk premium
                3) Bond ratings
        C) Liquidity
                1) Definition of liquidity
                2) Loans versus bonds
                3) Differences in liquidity across bonds
                        a) Introduction
                        b) Time it takes to sell
                        c) Costs of selling bonds
                              i) Introduction
                              ii) Brokers
                              iii) Dealers and the bid-ask spread
                4) Liquidity premium
        D) Taxes

IV) Interest rate differentials

A) Introduction

This part of the course is about something I briefly alluded to earlier and told you to forget. Consider two bonds that both promise to pay an IOU in September 2025. Earlier I told you to assume, for simplicity, that the yields on these two bonds had to be the same. Actually, not true. Assets which are supposed to pay off at exactly the same point in time *can* pay different yields or interest rates. Differences between yields or interest rates on bonds or loans that are supposed to pay off at exactly the same point in time are called interest rate (or yield) "differentials." We'll examine them here. You will see that interest rate differentials are the result of three things: differences across bonds and/or loans in the degree of default risk; differences in liquidity; and differences in the way the tax code treats income earned from this investment versus that investment.

B) Default risk

1) Definition of default risk

Default risk is the risk that a borrower will fail to make future payments as promised - will pay later than promised, or pay less than promised, or not pay at all. Any of these failures is called "default." A borrower who took a loan defaults by failing to make payments as promised in the loan contract. A borrower who issued a bond defaults by failing to pay the promised IOUs, or paying them later than promised.

Default happens a lot. When the borrower is a person, a business, or a municipality (city or county), default is usually part of a process called "bankruptcy" in which the borrower declares that it cannot pay its debts as promised. Bankruptcy involves lawyers and courts. We'll get back to bankruptcy later on. National or state govenments cannot declare bankruptcy, in the strict legal sense. But they sometimes fail to pay. The state of New York has never defaulted on bonds it issued, but the state of Pennsylvania has. Many countries have defaulted on their bonds at one time or another. Some countries, such as Argentina and Greece, have defaulted several times! Only a few countries, including America, Britain and Canada, have *never* defaulted.

2) Default risk premium

If investors perceive that there is a chance a bond issuer will default, they won't pay as much for the bond. The bond's price will be low, relative to its IOU(s). Remember the relationship between bond prices and yields: lower price, higher yield. Thus, default risk *raises* the yield on a bond. Generally, the higher the perceived probability of default on a bond, the higher its yield will be. Something similar goes for interest rates on loans: lenders demand higher interest rates on loans to borrowers with relatively high probability of default.

When one asset pays a higher return than another asset of the same maturity, the difference or "spread" between the higher-return asset and the lower-return asset is called a "premium." So if you are comparing two bonds of the same maturity, and bond A has higher default risk than bond B, the difference (spread) between the higher yield on bond A and the lower yield on bond B is called a "default risk premium." We would say that the relatively high yield on the default-riskier bond "contains" a default risk premium. The interest rate on a loan can also "contain" a default risk premium.

Frequently, people want to calculate a bond's default risk premium relative to a benchmark bond that is generally considered to have no default risk. Of course, to make such a comparison you must be careful to match the timing of the two bonds' payments - same maturity, zero-coupon bond versus coupon bond and so on. You wouldn't compare a ten-year zero-coupon bond with a five-year zero-coupon bond. Nor would you compare a ten-year zero coupon bond with a ten-year *coupon* bond.

For bonds that pay off in dollars, the no default-risk benchmark everyone uses is American Treasury bonds. Treasury bonds come in almost all maturities, ranging from one month to thirty years, as both coupon and zero-coupon bonds. So you can find a Treasury bond to match the payoff timing of almost any bond. American Treasury bonds are practically free of default risk - at least, that is what investors seem to think. Presumably the U.S. would default on its bonds in the event of a world-destroying event such as nuclear war, asteroid hit or zombie apocalypse. But very few investors attempt to prepare for that sort of thing.

One can describe the yield on a bond with default risk as being equal to the yield on an Amerrican Treasury bond of matching maturity plus a default risk premium. Here's an example. Let $_m i_t^{TREAS}$ denote today's yield on a zero-coupon Treasury bond paying off in m years. Let $_m i_t^{CORP}$ denote the yield on a zero-coupon corporate bond subject to default risk, also paying off in n years. Then:

$$_m i_t^{CORP} = {}_m i_t^{TREAS} + Default\ Risk\ Premium$$

3) Bond ratings

How do you know the default risk on a bond? You don't, really. It is a matter of opinion. To gauge the probability of default on a bond issued by a corporation, you would want to examine the corporation's finances, think about the future of its business, and so on. To gauge the probability of default by a country, you would think about politics.

There are, however, several companies that make a business of assessing the default risk on bonds sold in American financial markets, and publishing their assessments. These assessments aren't perfect. People in bond markets don't take them as gospel. But they are generally considered to be useful information.

One of these companies is called Moody's; another is called Standard & Poor's. These companies present their assessments, not as percent probabilities of default, but as letter grades, very much like the grades you get in school. A worse grade means a higher probability of default, and the best grades are A's! Standard & Poor's grade system is:
AAA means smallest probability of default
AA
A
BBB
BB

B
CCC
CC
C   means largest probability of default
D   really, bad; means the bond is actually currently in default.

Bonds with higher grades, from AAA down through BBB, are called "investment grade" bonds. Bonds with grades lower than BBB are called "junk" or "speculative grade" bonds.

A rating company rates a bond when it is issued, and may continue to monitor the issuer's financial condition after that, publishing new ratings of the bond when its assessment of the bond's default risk changes a lot.

How do bond rating companies make money on the operation? They are *paid* to rate a bond. Who pays the rating company to rate a bond? In most cases, the issuer of the bond. A bond issuer will pay to have its bond rated *even if the bond is sure to receive a low rating*. Why? We'll get back to this later.

C) Liquidity

1) Definition of liquidity

This is another reason for rate differentials, and an important concept in financial markets that comes up in many different contexts. We will examine liquidity fully in the next section of the course. Here, I will just introduce the concept.

In financial markets, "liquidity" refers to the amount of time, cost and trouble it takes to sell an asset. An asset is more liquid the easier/cheaper/faster it is to sell the asset. An asset is less liquid - more illiquid - the harder/more costly/longer it takes to sell the asset.

The concept of liquidity can be applied to things other than bonds and loans. One can talk about the liquidity of a house. As an asset, a house is less liquid than most stocks and bonds. It can take a long time to sell a house. There are also monetary costs. A lawyer must be paid to make sure that the seller's title to the house is free and clear, and to draw up the sales contract. Usually, the seller of a house engages a realtor and pays the realtor a fee to help sell the house. (Ask your parents about this.)

Cars are also pretty illiquid. Have you or your parents ever tried to sell a car? You have to put up a notice on Craigslist, show the car to potential buyers and so on. Selling a car this way is such a pain that many people sell their old cars to used car dealers. A dealer will buy your old car immediately. But selling a car to a used car dealer involves a hidden monetary cost. A used car dealer won't pay you as much for the car as you would get if you took the time and trouble to sell the car to another person on Craigslist. He buys the car in order to sell it to someone else (or sell it to another dealer who will sell the car to someone...) at a higher price, about equal to the price you could get yourself if you sold the car on Craigslist. From your point of view, the difference between the Craigslist price and the price the dealer pays you is *a cost of selling your car*. Essentially, the dealer's mark up is what you are paying the dealer to sell your old car for you.

What I said about the used car dealer will be important in a minute. If you weren't paying attention, read it again.

2) Loans versus bonds

Generally, loans are less liquid than bonds. To sell a bond, you just sell the piece of paper. To sell a loan, the original lender has to find someone willing to pay for the right to receive the loan payment(s) promised by the borrower. Then the borrower's contractual obligation must be transferred from the original lender to the buyer of the loan. This can involve costly legal work.

3) Differences in liquidity across bonds

a) Introduction

Though bonds are generally more liquid than loans, some bonds are more liquid than others. Some bonds can be sold quickly; some take longer to sell. More time to sell, less liquid (more illiquid). There are costs to selling bonds analogous to the realtor's fee to help sell a house and the dealer's mark up on a used car. Higher costs, less liquid (more illiquid).

b) Time it takes to sell

What do I mean by the amount of time it takes to sell a bond? After all, if you are willing to take a low enough price, you can sell anything fast. What I mean is, the time it takes to sell a bond for the highest price you are likely to get - the highest price that many potential buyers out there would be willing to pay for the bond.

For some bonds, it does not take long to sell the bond for such a price. For others it can take a long time to find the kind of buyer who values the bond relatively highly. In the next section of the course I will explain why.

c) Costs of selling bonds

i) Introduction

Remember what I said about realtors and used care dealers? Most people who sell bonds don't handle the sale themselves. They use the services of bond brokers or bond dealers, who must be paid.

ii) Brokers

A bond broker is analogous to a realtor in the house market. A bond broker does not buy your bond himself. He finds a buyer and arranges the sale of the bond from you to the buyer. For this service the broker charges a fee (or "commission"). Brokers can charge different fees to sell different types of bonds (different maturities, different issuers). The higher the fee a broker charges to sell a bond of a given type (maturity and issuer), the less liquid is that type of bond.

iii) Dealers and the bid-ask spread

A bond dealer is a like a used car dealer. She buys your bond from you. She does not want to hold the bond herself; her goal is to sell the bond on to someone else at a higher price. The price a bond dealer pays for a bond of a given type - maturity and issuer - is called the "bid" price. The higher price the dealer charges when she sells the same type of bond is called the "ask" price. The dealer makes money on the difference between the ask and bid prices. From your point of view, this price difference is *a cost of selling your bond*. Essentially, you are paying the dealer to sell your bond for you; the price difference is what you are paying the dealer to do this. Like fees charged by brokers, the difference between a dealer's bid and ask prices can vary cross bonds (maturity and issuer). The larger is the dealer's price difference for a bond of a given type, the less liquid (or more illiquid) is that type of bond.

Remember that when we talk about bonds we usually don't talk about bond prices directly; we talk about bonds' yields. This is true when we talk about the cost of selling bonds through a dealer. We don't talk about the difference between the bid and ask *prices*. Instead we talk about the **"bid-ask spread"** in *yields*. Because the bid price is lower than the ask price, the yield calculated from the bid price is *higher* than the yield calculated from the ask price. The "bid-ask spread" for a bond is the relatively high yield calculated from dealers' (low) bid price *minus* the relatively low yield calculated from dealers' (higher) ask price. The larger is dealers' bid-ask spread for a bond of a given type, the less liquid is that type of bond.

4) Liquidity premium

Liquidity is a good thing from an investor's (lender's) point of view. Thus, the less liquid a bond is, the lower is the price that people are willing to pay for the bond, relative to its IOU(s). Relatively illiquid bonds have relatively high yields. (Remember, when I say a bond "less liquid" I mean it would take a relatively long time to sell the bond at a "high" price - as defined above, or brokers charge a relatively high fee to sell the bond, or dealers' bid-ask spread is relatively large for the bond.)

The difference (spread) between the higher yield on a less-liquid bond and the lower yield on a more liquid bond of the same maturity and type (zero-coupon versus coupon) is called a "liquidity premium." We would say that the yield on the less-liquid bond "contains" a liquidity premium.

The language here is screwy. A premium in a bond's yield - think of a default-risk premium - reflects something *bad* about the bond. Here the bad thing is *illiquidity*. So liquidity premium *ought* to be called an "illiquidity premium." But it isn't. The English language has many inconsistencies. There are many better languages: Latin of course; also Italian and Spanish. Sadly, we must take the world as we find it.

Frequently, people want to calculate a bond's liquidity premium relative to a benchmark bond that is as liquid as a bond can be. For bonds that pay off in dollars, the perfectly-liquid benchmark is, once again, American Treasury bonds. No bonds can be sold more quickly than Treasury bonds. Brokers' fees and dealers' bid-ask spreads are smallest for Treasury bonds.

As in the case of default-risk premia, one can describe the yield on a bond less liquid than a Treasury bond as being equal to the yield on a Treasury bond of matching maturity *plus* a liquidity premium. What if a bond is less liquid than a Treasury bond *and* has default risk? Then we would say that its yield contains *both* a liquidity premium *and* a default-risk premium. Let $_m i_t^{TREAS}$ denote today's yield on a zero-coupon Treasury bond paying off in *n* years. Let $_m i_t^{CORP}$ denote the yield on a zero-coupon corporate bond, also paying off in *n* years, which is less liquid than a Treasury bond *and* subject to default risk. Then:

$$_m i_t^{CORP} = {_m i_t^{TREAS}} + Default\ Risk\ Premium + Liquidity\ Premium$$

Finally, what about loans? Loans are generally less liquid than bonds. That is a bad thing. So lenders will usually require a higher interest rate on a loan than the yield on a bond subject to the same degree of default risk. The difference between the interest rate on a loan and the yield on a bond with the same degree of default risk is a liquidity premium.

D) Taxes

I guess you're getting the idea now: if an asset is relatively good in some way - low default risk, high liquidity, smells good or whatever - it will have a lower yield or interest rate.

The peculiarities of the tax code create interest rate differentials on certain types of coupon bonds. Coupon payments on bonds issued by corporations or by the American Treasury count as income for purposes of Federal income tax. Holding such bonds raises your income tax bill. Coupon payments on bonds issued by states and municipalities do *not* count as income for Federal income tax. Coupon payments on these bonds are (income) tax-free. That is a good thing! It tends to rise the prices of state and municipal bonds, lowering their yields. For this reason, yields on state and muncipal bonds with low default risk can be *lower* than yields on Treasury bonds of matching maturity.

©**Christopher Hanes   Economics 450   Class notes**

V) Term structure
  A) Introduction
    1) Definitions: term structure, yield curve
    2) A theory of term structure
  B) Review: "expected value" of a variable
    1) Definition
    2) Examples
  C) Expectations hypothesis of the term structure
    1) Introduction
    2) Assumption; people care only about expected values of returns to possible investments
    3) Examples
      a) A bond with default risk
      b) Expected value of the rate of return to selling a bond before it matures
      c) Expected value of rolling over a series of one-year bonds or loans
      d) Expected value of rolling over a series of overnight loans
    4) The yield curve
      a) Introduction
      b) Situations and resulting yield curve
        i) Expect future overnight rates may be higher, lower, or same
        ii) Expect overnight rates may rise
        iii) Expect overnight rates may fall
        v) Summary
  D) Term premiums
    1) A puzzle
    2) Solution: term premiums
    3) Explanation of term premiums: interest-rate risk

V) Term structure

A) Introduction

**1) Definitions: term structure, yield curve**

This part of the course is about differences in yields or interest rates across bonds or loans that are due to differences in the time they pay off, for example, the difference between the yield on a bond that is supposed to pay its IOU in September 2025 and the yield on a bond that pays off in September 2026. The pattern of differences in yields and interest rates associated with maturity are called "term structure" - "term" is another word for maturity. Bonds of long maturity are called "long-term" bonds; bonds of short maturity are called "short-term" bonds, and so on.

Of course, as you learned in the last section, there are also differentials in yields or interest rates across bonds and loans that pay off at the *same* point in time due to differences in default risk, liquidity and tax treatment. To focus on term structure *per se*, not mix it up with those other interest-rate differentials, people compare yields on a set of bonds that pay off at different points in time but are the same in terms of default risk, liquidity and tax treatment. The easiest way to do this is to use a set of bonds that are all from the same issuer. And in the U.S. the easiest way to get a set of bonds of many different maturities, all from the same issuer, is to use U.S. Treasury bonds. Here is a list, taken from a financial website, of yields on single-payment Treasury bonds of various maturities ("bills" and "zero-coupon" bonds), on August 3 2020:

Date	1 mo	2 mo	3 mo	6 mo	1 yr	2 yr	3 yr	5 yr	7 yr	10 yr	20 yr	30 yr
08/03/20	0.09	0.09	0.10	0.11	0.12	0.11	0.13	0.22	0.40	0.56	1.01	1.23

This pattern would be called "the term structure of interest rates" on August 3, 2020.

1

Frequently, people visualize term structure by making a graph called a "yield curve." To do this you get a list of today's yields on a set of bonds that pay off at different points in time but are otherwise exactly the same, like the list above. You make a graph of these numbers with the maturity of a bond on the horizontal axis and the corresponding bond's yield on the vertical axis. In the U.S., when people talk about "the yield curve" they nearly always mean a graph made of yields on Treasury bonds - the "Treasury yield curve." A yield curve made from the above numbers is:
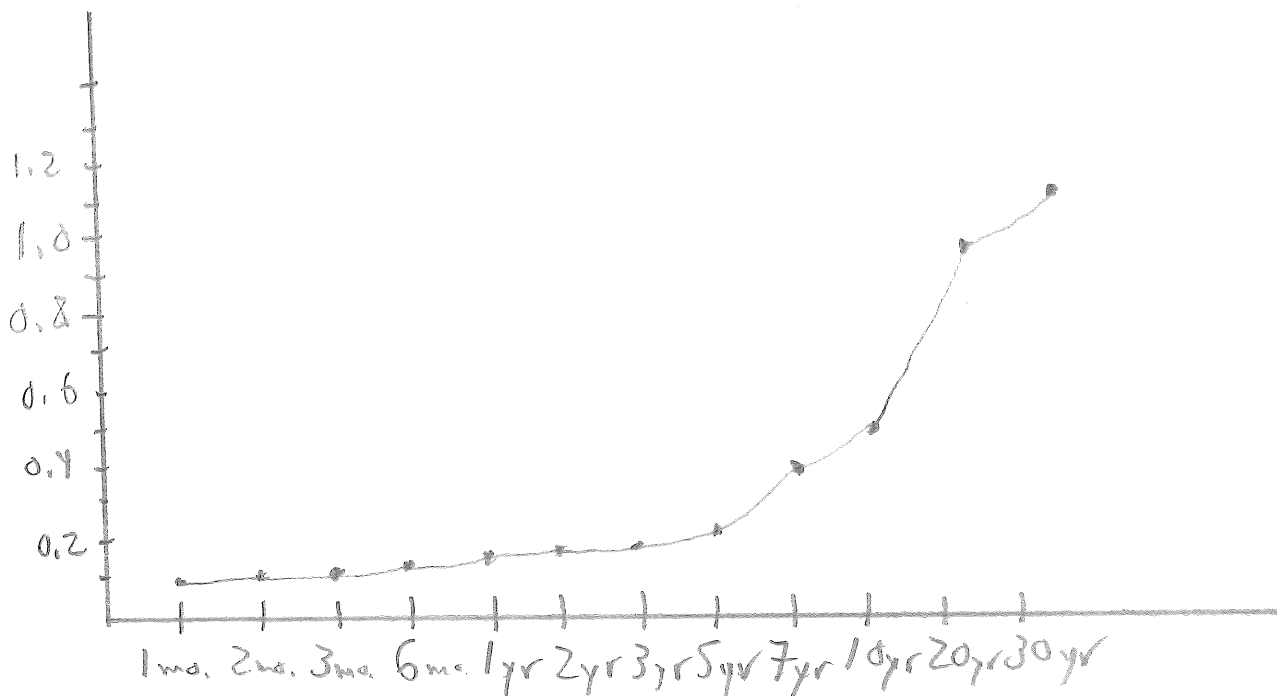
You can see that on August 3, 2020, the yield curve was "upward sloping": yields on longer-maturity bonds tended to be higher than yields on shorter-maturity bonds. That is true on many days, but it is not *always* true. On some days you will find that the yield curve is downward-sloping (yields on longer-term bonds tends to be lower than yields on shorter-term bonds) or flat (yields on bonds of all maturities are about the same).

2) A theory of term structure

There is a theory that explains the shapes of yield curves - why they are upward-sloping, downward-sloping or flat on a particular day. This theory is believed by just about everybody, certainly everybody who works for the Federal Reserve system, and as far as I know everybody who works in finance. It starts with a basic, simple theory of term structure and then adds a complication.

The basic, simple theory is called the "expectations hypothesis of the term structure." This is the answer to the question, "what would the term structure be if investors cared only about the *expected values* of the returns on their investments?" Here "expected value" means a concept from statistics, the "expected value of a variable." We'll begin this section by reviewing that concept. The expectations hypothesis of the term structure can explain why the yield curve is sometimes upward-sloping, sometimes downward-sloping, sometimes flat.

But we will find that the expectations hypothesis is inconsistent with a well-know fact about the yield curve. To account for that fact we have to add a complication to the expectations hypothesis. The complication is called "term premiums." And the notion of term premiums is related to something we covered earlier: interest-rate risk (or duration risk).

B) Review: expected value of a variable

1) Definition

Remember from stats: a "random variable" is a variable for which you do not know the value, but you have an idea of the possible values the variable *could* take and the probabilities that it could take each of those values. You might be talking about a variable that hasn't occurred yet, e.g. the temperature outside at noon tomorrow. Or you might be talking about a variable which has already occurred, but which you haven't observed yet, e.g. the temperature outside right now but you haven't been outside yet. Either way, it is a random variable.

The expected value of a variable is one number that roughly summarizes your beliefs about the possible values that variable might take and the associated probabilities. To calculate the expected value of a random variable, you could make a table that shows, in one column, all the possible values the variable could take, and in another column the probabilities of each possible value. Then you multiple each possible value by its probability. Then you add up all of those multiples to get one number. That number is the "expected value" of the random variable. An important thing about the expected value of a variable, on this statistical definition, is that it is *not* necessarily the same as "the value you expect the variable to take."

2) Examples

Here are some examples.

First example: flipping a fair coin and giving you a prize. Suppose I am going to flip a fair coin and give you $4 if it comes up heads, $8 if it comes up tails. Table:

Possible scenarios	Probability	Prize in that case
Heads	1/2 or 0.50 or 50 percent	$4
Tails	1/2 or 0.50 or 50 percent	$8

$$E[Prize] = \frac{1}{2} \bullet 4 + \frac{1}{2} \bullet 8 = 2 + 4 = 6$$

The expected value of your prize is $6. That is not the value you expect your prize to take! Your prize cannot possibly be $6. Your prize will be either $4 or $8.

Second example: flipping a trick coin and giving you a prize. In this case the coin is weighted so that it is more likely to come up heads. Table:

Possible scenarios	Probability	Prize in that case
Heads	1/4 or 0.25 or 25 percent	$4
Tails	3/4 or 0.75 or 75 percent	$8

$$E[Prize] = \frac{1}{4} \bullet 4 + \frac{3}{4} \bullet 8 = 1 + 6 = 7$$

Here the expected value of your prize is $7, even though your prize cannot possibly be $7. Your prize will be either $4 or $8.

Third example: flipping a magical coin. A wizard has a magical coin which can land on its edge and stay there. If the coin comes up heads, he'll give you $6; tails, $9. If it lands on its edge, the wizard will give you $12. Table:

Possible scenarios	Probability	Prize in that case
Heads	1/3 or 0.333.. or 33.3... percent	$6
Tails	1/3 or 0.333.. or 33.3... percent	$9
Edge	1/3 or 0.333.. or 33.3... percent	$12

$$E[Prize] = \frac{1}{3} \bullet 6 + \frac{1}{3} \bullet 9 + \frac{1}{3} \bullet 12 = 2 + 3 + 4 = 9$$

Here, the expected value of the prize, $9, *is* one of the actual values the price can take.

What is the expected value of a variable if you know, for sure, the value that the variable will take? It is just that known value. To continue the examples above, if you know for sure I will give you $5, the expected value of your prize is $5.

C) Expectations hypothesis of the term structure

1) Introduction

The "expectations hypothesis of the term structure" is a model that will be one component of our complete theory of the term structure. It is part of the truth, but not the whole truth. For the expectations-hypothesis model, we assume that *investors care only about the expected values of the future returns to their investments*. We figure out what the term structure would look like if that assumption were true. That is the expectations hypothesis of the term structure. Don't forget: the assumption is *not* true. But going through this exercise helps us think about things.

2) Assumption: people care only about expected values of return to possible investments

What do I mean when I say "investors care only about the expected values of returns on their investments"?

Consider an investor - a lender, in our terms - who is willing to give up money today in exchange for a promise of money to be received at a specified future point in time. As an example, say the specified point in time is five years from now. There are many possible investments to consider - many ways to exchange money today for a promise of payment of money five years from now.

In the category of loans, the lender could make a loan that matures in five years. Or she could "roll over" a series of five one-year loans. Or make one two-year loan, rolling over into a three-year loan. And so on.

In the category of bonds, the lender could buy a Treasury bond that matures in five years. Or she could buy another kind of bond that matures in five years. Or she could buy a longer-term bond (a bond that mature in more than five years) and sell it after holding it just five years. Or she could buy a short-term bond that will mature in just one year, planning to "roll over" the short-term bond into a series of four more one-year bonds, cashing out five years from now. And so on.

The payment that will result from a five-year loan with no default risk, or a five-year bond with no default risk, is *certain*. But the payment that will result from any of the other investments is *uncertain*. A five-year loan or bond with default risk may or may not pay off. The price received for a longer-term bond sold five years from now will depend on what bond yields are at that time. The return from roling over short-term loans or short-term bonds depends on what short-term interest rates or short-term bond yields turn out to be in the future.

How does an lender compare an investment that has a certain payoff with an investment that has an uncertain payoff? Or two investments that both have uncertain payoffs? For the expectations hypothesis, we assume that the lender thinks about the *expected value* of the return to each possible investment, and buys the one that has the highest expected value.

Under this assumption, there is just one possible equilibrium in financial markets: bond prices and interest rates on loans must adjust to *equalize* the expected values of the returns to all investments that will pay off (or possibly pay off - remember default risk) at a given future point in time.

To show you what this means and how it works, I'll go through some examples. You'll see that the assumption has implications for many things other than term structure. But we'll get to term structure eventually. It will follow from our last example.

3) Examples

a) A bond with default risk

Consider a single-payment bond that has default risk. You know that the yield on that bond must be higher than the yield on a bond with no default risk, i.e. a Treasury bond - that's the default-risk premium. That is the same as saying that the price of the bond must be lower, relative to its IOU, than the price of a Treasury bond of matching maturity. But *exactly how much lower* is the price, *exactly how much higher* is the yield, of the risky bond relative to the Treasury bond? The assumption that we make for the expectations hypothesis determines those things exactly. The equilibrium market price of this bond must equalize the expected value of the return to holding the bond, accounting for the risk of default, to the yield on a matching-maturity Treasury bond.

To start, let's make a table that describes a hypothetical investor's beliefs about the returns she might receive if she buys the risky bond and calculate the expected value of the return. As in our examples of coin flips and prizes above, we think in terms of scenarios, probabilities of each scenario, and the "prize" you get under each scenario.
In the case of the default-risky bond there are two scenarios: default, and no default. Let's say that people in financial markets estimate the probability of default on the bond to be x, where x is a number such as 1/10 (a 10 percent probability of default), 1/4 (a 25 percent probability of default), or 1/2 (a 50 percent probability of default). That means, of course, that the probability of *no* default is *(1-x)*. Here is our table so far:

Possible scenarios	Probability	Realized return
Default	x	
No default	(1-x)	

To complete the table, we need to figure out what the realized return will be in each scenario.
If there is no default and the bond pays off as promised, the realized return on the bond is simply its yield:

$$_m i = \sqrt[m]{\frac{IOU}{P}} - 1 = \left(\frac{IOU}{P}\right)^{1/m} - 1$$

If the issuer of the bond defaults and pays nothing at all, the payoff on the bond will be zero. Thus in this event the realized return on the bond is:

$$_m i = \sqrt[m]{\frac{0}{P}} - 1 = \left(\frac{0}{P}\right)^{1/m} - 1 = -1 \text{ or } -100 \text{ percent.}$$

So our complete table is:

Possible scenarios	Probability	Realized return
Default	x	-1
No default	(1-x)	$\left(\frac{IOU}{P}\right)^{1/m} - 1$

5

The expected value of the return you will get from this bond, in terms of a fraction (before converting into percent) is:

$$E[Return] = x \bullet (-1) + (1-x) \bullet \left[\left(\frac{IOU}{P} \right)^{1/m} - 1 \right]$$

In equilibrium this expected return must be equal to the yield on an m-year Treasury bond:

$$_m i \; = \; x \bullet (-1) + (1-x) \bullet \left[\left(\frac{IOU}{P} \right)^{1/m} - 1 \right]$$

or, in terms of the risky bond's yield,

$$_m i \; = \; x \bullet (-1) + (1-x) \bullet \; _m i_{RISKYBOND}$$

How does the market get to this equilibrium?

Suppose that we were out of equilibrium, in such a way that the expected return to holding the risky bond were *lower* than the yield on a matching-maturity Treasury bond. That would make the risky bond a bad deal. No one would want to buy it. Anyone holding it would want to sell it. So the market price of the risky bond must fall. As the bond's price falls, the expected return to holding it increases. (Look up at the expression that defines $E[Return]$, can you see how that is true?) The price must continue to fall until the two sides of the equation are equal.

If the bond's price is too low - if the expected return to holding the risky bond is higher than the Treasury bond yield - the opposite happens.

In bond markets, at least American bond markets nowadays, equilibrium happens very fast. In this class, we'll just assume the equilibrium always holds immediately.

So, the market price of the bond in question will be the price that solves the above equation. And that determines the yield on the risky bond.

For a specific example, let's say that we're talking about a one-year single-payment bond issued by the IBM Corporation. The IOU on the bond is $1000. People in financial markets estimate the probability of default on the bond to be 1/4. And the yield on one-year zero-coupon Treasury bonds is five percent.

What is the equilibrium market price of the IBM bond? It is the price that makes this true:

$$0.05 \; = \; \frac{1}{4} \bullet (-1) + \frac{3}{4} \bullet \left[\frac{1000}{P_{IBM}} - 1 \right]$$

This is an algebra problem. Take the equation and solve for P_{IBM}. Your answer should be:

$$P_{IBM} \; = \; 1000 \; \frac{3}{4} \; \frac{1}{1.05} \; = \$714.29$$

b) Expected value of the rate of return to selling a bond before it matures

Remember the rate of return to selling a bond before it matures, denoting the buying price as $_m P_t$ and the selling price as $_{m-n} P_{t+n}$:

$$RER = \sqrt[n]{\frac{_{m-n}P_{t+n}}{_mP_t}} - 1 = \left(\frac{_{m-n}P_{t+n}}{_mP_t}\right)^{1/n} - 1$$

$_{m-n}P_{t+n}$ is uncertain as of time t, when the bond is purchased. Thus, at that time the rate of return is a random variable. The expected value of the rate of return is denoted:

$$E[RER] = E\left[\sqrt[n]{\frac{_{m-n}P_{t+n}}{_mP_t}} - 1\right] = E\left[\left(\frac{_{m-n}P_{t+n}}{_mP_t}\right)^{1/n} - 1\right]$$

To think about this, an investor maps out possible future scenarios with respect to the future price of the bond, which is to say possible future secanrios with respect to bond yields in the future: the bond's future price will be determined by bond yields prevailing at the time the bond is sold.

For a simple example, let's say that our bond market participants think there are only three possible scenarios with respect to future bond yields. Call them A, B and C. The probability of scenario A is x_A. The probability of scenario B is x_B. That means, of course, that the probability of scenario C is $(1 - x_A - x_B)$. To keep things simple, let's skip a step and say that our bond market participants have already done the math for this particular bond, with its particular IOU, to get the price of the bond under each of the three possible yield scenarios. Under scenario A the bond's price will be is P_A. The bond's price under scenario B will be P_B ; under scenario C, P_C. Which gives us this table:

Possible scenarios	Probability	Realized return
A	x_A	$\left(\frac{_{m-n}P_{At+n}}{_mP_t}\right)^{1/n} - 1$
B	x_B	$\left(\frac{_{m-n}P_{Bt+n}}{_mP_t}\right)^{1/n} - 1$
C	$(1 - x_A - x_B)$	$\left(\frac{_{m-n}P_{Ct+n}}{_mP_t}\right)^{1/n} - 1$

and:

$$E[RER] = x_A \bullet \left[\left(\frac{_{m-n}P_{At+n}}{_mP_t}\right)^{1/n} - 1\right] + x_B \bullet \left[\left(\frac{_{m-n}P_{Bt+n}}{_mP_t}\right)^{1/n} - 1\right] + (1 - x_A - x_B) \bullet \left[\left(\frac{_{m-n}P_{Ct+n}}{_mP_t}\right)^{1/n} - 1\right]$$

The assumption for the expectations hypothesis implies that, in equilibrium, the expected value of the rate of return to selling a bond n years from now (before the bond matures) must equal today's yield on a zero-coupon Treasury bond that matures n years from now. So it must be true that:

$$_ni_t = x_A \bullet \left[\left(\frac{_{m-n}P_{At+n}}{_mP_t}\right)^{1/n} - 1\right] + x_B \bullet \left[\left(\frac{_{m-n}P_{Bt+n}}{_mP_t}\right)^{1/n} - 1\right] + (1 - x_A - x_B) \bullet \left[\left(\frac{_{m-n}P_{Ct+n}}{_mP_t}\right)^{1/n} - 1\right]$$

The equilibrium current price of the bond $_mP_t$ is the value that solves this equation. The equation above is messy, but the equation becomes simple if you assume that the bond will be sold in just one year ($n=1$). Make up an example like that for yourself and solve for the price.

V

4) Expected value of the return to rolling over a series of one-year bonds or loans

One way a lender can arrange to get a payment coming n years in the future is by making a series of one-year loans, rolling over the principal and interest on the loans until cashing out at time $t+n$.

Another way is to buy one-year bonds, rolling over the IOU's on the bonds into more one-year bonds, until cashing out at time $t+n$.

Recall that we can describe both of these operations the same way. The lender starts out with $\$X$ (lending out that much, or spending that much on bonds). The amount of money he will have at the end when he cashes out is:

$$X(1+{}_1i_t)(1+{}_1i_{t+1})(1+{}_1i_{t+2})..........(1+{}_1i_{t+n-1})$$

Here I am using ${}_1i$ to denote the yield on a one-year bond or the interest rate on a one-year loan. The realized return on this operation, expressed on an annual basis as we express other returns, is:

$$\left(\frac{X(1+{}_1i_t)(1+{}_1i_{t+1})..........(1+{}_1i_{t+n-1})}{X}\right)^{1/n} - 1 = \left((1+{}_1i_t)(1+{}_1i_{t+1})..........(1+{}_1i_{t+n-1})\right)^{1/n} - 1$$

Future one-year yields or interest rates are uncertain . Thus, as of time t, when the operation begins, the return is a random variable. The expected value of the return is denoted:

$$E\left[\left((1+{}_1i_t)(1+{}_1i_{t+1})..........(1+{}_1i_{t+n-1})\right)^{1/n}\right] - 1$$

The assumption for the expectations hypothesis implies that, in equilibrium, the expected value of the return to rolling over one-year bonds or loans for n years must equal today's yield on a zero-coupon Treasury bond that matures n years from now. So:

$$_ni_t = E\left[\left((1+{}_1i_t)(1+{}_1i_{t+1})..........(1+{}_1i_{t+n-1})\right)^{1/n}\right] - 1$$

Well, this looks like a mess! How are we going to deal with all of this multiplication and taking nth roots? I CAN'T DEAL WITH THAT!

Relax. There is a handy approximation that makes all this *easy* to deal with. It is that the annualized rate of return to rolling over a series of one-year investments between now and n years from now is *approximately* - not exactly, but very close - equal to the average of the one-year interest rate or yield between now and n years from now. I mean this:

$$\left((1+{}_1i_t)(1+{}_1i_{t+1})..........(1+{}_1i_{t+n-1})\right)^{1/n} - 1 \approx \frac{{}_1i_t + {}_1i_{t+1} + {}_1i_{t+2} + ...+ {}_1i_{t+n-1}}{n}$$

The squiggly "equals sign" is a mathematical symbol used to denote "approximately equal." So, for expected value,

$$E\left[\left((1+{}_1i_t)(1+{}_1i_{t+1})..........(1+{}_1i_{t+n-1})\right)^{1/n}\right] \approx E\left[\frac{{}_1i_t + {}_1i_{t+1} + {}_1i_{t+2} + ...+ {}_1i_{t+n-1}}{n}\right]$$

Under the assumption for the expectations hypothesis,

$$_ni_t \approx E\left[\frac{{}_1i_t + {}_1i_{t+1} + {}_1i_{t+2} + ...+ {}_1i_{t+n-1}}{n}\right]$$

In this class, we will forget that this is an approximation and treat is as an ordinary equation.

8

5) Expected value of the return to rolling over a series of overnight loans

Here is our last example. It is the one that is going to lead into our theory of term structure and the yield curve. It is about "overnight loans." I haven't talked about overnight loans yet. To keep things simple, the shortest-maturity investments I have dealt with have been one-year bonds or loans (though I have mentioned that there are shorter-term bonds called "bills"). But in reality there is lots and lots of lending at the overnight maturity. This is the shortest maturity at which lending regularly takes place in America.

It literally means overnight: the lender gives the borrower money by the end of the business day, five or six pm; the borrower promises to repay the lender, with interest, at the beginning of the following business day. Over weekends, the lender gives the borrower money at the end of Friday, promising to repay on Monday morning. On weekends when financial markets will be closed Monday because of a national holiday, the borrower repays Tuesday morning. Across Good Friday, when financial markets are closed even though it's not a national holiday..well, let's forget about all that. In this class, we'll pretend there are no weekends or holidays, so that overnight lending takes place 365 days a year.

Fortunately, we needn't worry about overnight bonds. An overnight bond would be one that is issued on Tuesday, say, with the IOU supposed to be paid on Wednesday. There is no such thing!

We'll need to introduce more notation to deal with overnight loans. First, we'll say that the matury of the loan is "0," so that the interest rate on an overnight loan is $_0i$. Second, we'll denote today, tomorrow and so on by $d, d+1, d+2...$.Thus:

$_0i_d$ is the overnight interest rate today.

$_0i_{d+1}$ is the overnight interest rate tomorrow.

$_0i_{d+2}$ is the overnight rate day after tomorrow

　　　and so on.

You can "roll over" overnight loans just as you roll over one-year loans or bonds. This is done *a lot*. So we need to think about the annualized return to rolling over overnight loans for many days, cashing out at some point in the future.

Fortunately, the approximation I introduced just above, about the annualized return to rolling over a series of one-year loans and bonds, works for overnight loans too. Consider the annualized return to rolling over overnight loans starting today, ending a year from today. This is approximately equal to:

$$\frac{_0i_d + _0i_{d+1} + _0i_{d+2} + ... + _0i_{d+365-1}}{365}$$ (Remember, we are pretending there are no weekends or holidays.)

If we roll over overnight loans from today until *two* years from today, the return is approximately:

$$\frac{_0i_d + _0i_{d+1} + _0i_{d+2} + ... + _0i_{d+729}}{730}$$

And so on.

Future overnight loan rates are uncertain. Thus, as of today, when the operation begins, these returns are random variables.
The assumption for the expectations hypothesis implies that, in equilibrium, the expected value of the return to rolling over overnight loans for one year must equal today's yield on a one-year zero-coupon Treasury bond.

$$_1i_t = E\left[\frac{_0i_d + _0i_{d+1} + _0i_{d+2} + ... + _0i_{d+365-1}}{365} \right]$$

Note that I am mixing up the different frequencies of my time indicators here - mixing up the t's with the d's. Sorry, can't avoid that.

Going on, the expected value of the return to rolling over overnight loans for *two* years must equal today's yield on a *two-year* zero-coupon Treasury bond:

$$_2i_t = E\left[\frac{_0i_d + _0i_{d+1} + _0i_{d+2} + \dots\dots + _0i_{d+729}}{730}\right]$$

And so on. This is how we will get term structure and the yield curve.

4) The yield curve

a) Introduction

Well, at long last here we are: the term structure. The expectations hypothesis explains it in terms of the last example we went over, the one about the relationship between bond yields and expected future overnight rates. The idea is that expectations of future overnight rates determine today's yields on Treasury bonds: the yield on an m-year Treasury bond is equal to the expected value of the average value of the overnight rate from today through m years from now. All other yields and interest rates are equal to yields on Treasury bonds at matching maturities *plus* default-risk and liquidity premiums. (Recall that the assumption for the expectations hypothesis also tells you exactly how big a default-risk premium should be, given the probability of default estimated by financial market participants.) The shape of the yield curve - whether it is flat, upward or downward-sloping - is just a reflection of expectations of what will happen to overnight rates in the future.

You might be asking yourself: "what is so special about overnight rates? Why are they the starting point? Why don't we say that Treasury yields determine expected overnight rates, rather than the other way around? The equations in part D) would work the same either way." Good question! The answer is that we believe that in a special way overnight rates are "*exogenous*" to financial markets, determined by factors *outside* financial markets. Overnight rates are determined by the policymakers in the central bank - in America, the members of the Federal Reserve's "Federal Open Market Committee" (FOMC). We'll return to this point below.

Right now, we're going to further examine the relationship between Treasury yields and expected future overnight rates under the expectations hypothesis. We will figure out the conditions under which the yield curve would be flat *versus* upward-sloping or downward-sloping. We'll see that the yield curve is:

- *upward-sloping* if market participants believe future overnight rates are likely to be higher in the future than they are today. That is, future overnight rates might be about the same as they are today or higher, but it is very unlikely they will be lower than current overnight rates.

- *downward-sloping* if market participants believe future overnight rates are likely to be lower in the future than they are today. Future overnight rates might be about the same as they are today or lower, but it is very unlikely they will be higher than current overnight rates.

- *flat* if market participants believe future overnight rates may be the same as they are today, or lower, or high - no reason to think they will go in a particular direction.

I'll demonstrate this with examples. For our examples, our yield curve graphs will show today's overnight rate $_0i$ along with today's Treasury bond yields at the one, two, and three-year maturities ($_1i$, $_2i$, $_{.2}i$). Real yield curves, the one you find in newspapers and on financial websites, show Treasury yields at maturities from three months to thirty years.

b) Beliefs about future overnight rates and resulting yield curve

i) Future overnight rates may rise or stay the same

Suppose that the overnight rate today is 5 percent. People think that it will remain around 5 percent for six months. On any particular day within that period the overnight rate may be a bit higher or lower than five percent, but it will be five percent on average across the six months. After that, the rate may rise or stay the same. To keep things simple, say a probability of 1/2 that the rate will rise. If it rises, it will rise by one percent to 6 percent.

The first point on the yield curve is easy: $_0i$ is 5 percent.

What is $_1i$? It is the expected value of the average value of the overnight rate from today through one year from today.

Scenario	Probability	Average value of $_0i$
A (rise)	1/2	$5\frac{1}{2}$ (six months of 5 percent, six months of 6 percent)
B (stay same)	1/2	5

$$_1i = \frac{1}{2} \bullet 5\frac{1}{2} + \frac{1}{2} \bullet 5 = 5\frac{1}{4} = 5.25$$

And $_2i$? That's the expected value of the average value of the overnight rate from today through *two* years from today.

| Scenario	Probability	Average value of $_0i$
A (rise)	1/2	$5\frac{3}{4}$ (six months of 5 percent, eighteen months of 6 percent)
B (stay same)	1/2	5

$$_2i = \frac{1}{2} \bullet 5\frac{3}{4} + \frac{1}{2} \bullet 5 = 5\frac{3}{8} = 5.375$$

Finally, $_3i$ is the expected value of the average value of the overnight rate from today through three years from today.

Scenario	Probability	Average value of $_0i$
A (rise)	1/3	$5\frac{5}{6}$ (six months of 5 percent, thirty months of 6 percent)
B (stay same)	1/3	5

$$_3i = \frac{1}{2} \bullet 5\frac{5}{6} + \frac{1}{2} \bullet 5 = 5\frac{5}{12} = 5.416$$

Putting all this together on the graph, the yield curve is upward sloping:
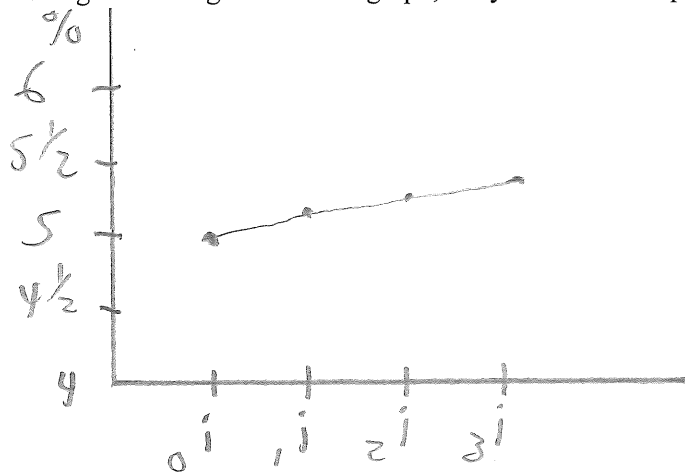
The yield curve is *generally* upward-sloping when market participants believe future overnight rates are likely to be higher in the future than they are today, or the same, but unlikely to be lower. Make up some examples for yourself.

ii) Future overnight rates may fall or stay the same

In this case, as before, the overnight rate today is 5 percent and people think that it will remain around 5 percent for six months. But after that, the rate may *fall* or stay the same, with probability of 1/2 for each. If the rate falls, it will fall by one percent to 4 percent. In this case, as in the previous one, the first point on the yield curve (current $_0i$) is 5 percent.

$_1i$ is the expected value of the average value of the overnight rate from today through one year from today.

Scenario	Probability	Average value of $_0i$
A (fall)	1/2	$4\frac{1}{2}$ (six months of 5 percent, six months of 4 percent)
B (stay same)	1/2	5

$$_1i = \frac{1}{2} \bullet 4\frac{1}{2} + \frac{1}{2} \bullet 5 = 4\frac{3}{4} = 4.75$$

$_2i$ is the expected value of the average value of the overnight rate from today through two years from today.

| Scenario	Probability	Average value of $_0i$
A (fall)	1/2	$4\frac{1}{4}$ (six months of 5 percent, eighteen months of 4 percent)
B (stay same)	1/2	5

$$_2i = \frac{1}{2} \bullet 4\frac{1}{4} + \frac{1}{2} \bullet 5 = 4\frac{5}{8} = 4.625$$

$_3i$ is the expected value of the average value of the overnight rate from today through three years from today.

Scenario	Probability	Average value of $_0i$
A (fall)	1/2	$4\frac{1}{6}$ (six months of 5 percent, thirty months of 4 percent)
B (stay same)	1/2	5

$$_3i = \frac{1}{3} \bullet 4\frac{1}{6} + \frac{1}{2} \bullet 5 = 4\frac{7}{12} = 4.583$$

so the yield curve is downward-sloping:
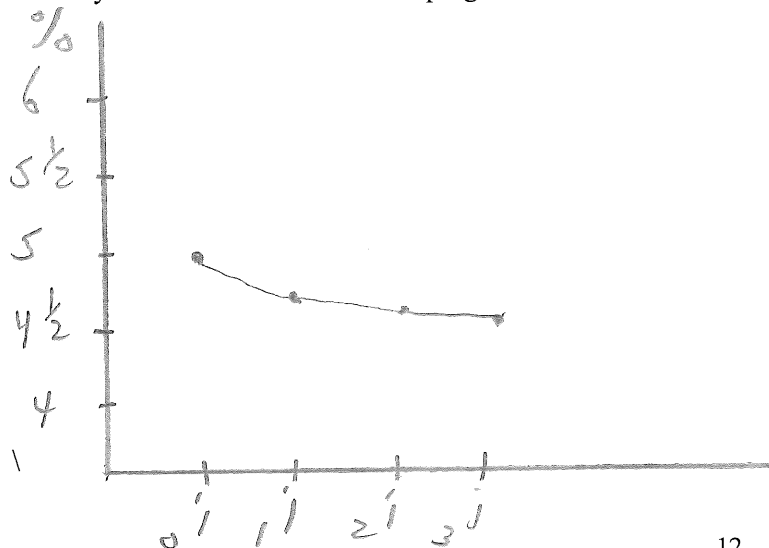
The yield curve is *generally* downward-sloping erally true when market participants believe future overnight rates are likely to be lower in the future than they are today, or the same, but unlikely to be higher.

iii) Future overnight rates may rise, fall, or stay the same

In this case the overnight rate today is 5 percent. People think that it will remain around 5 percent for about six months. After that, the rate may rise, fall, or or stay the same - all equally likely, with a probability of 1/3 each. If the rate rises, it will rise one percent, to 6 percent. If it falls, it will fall by the same amount, to 4 percent. Again, first point ($_0i$) is 5%.

$_1i$ is the expected value of the average value of the overnight rate from today through one year from today.

Scenario	Probability	Average value of $_0i$
A (rise)	1/3	$5\frac{1}{2}$ (six months of 5 percent, six months of 6 percent)
B (fall)	1/3	$4\frac{1}{2}$ (six months of 5 percent, six months of 4 percent)
C (stay same)	1/3	5

$$_1i = \frac{1}{3} \bullet 5\tfrac{1}{2} + \frac{1}{3} \bullet 4\tfrac{1}{2} + \frac{1}{3} \bullet 5 = 5$$

$_2i$ is the expected value of the average value of the overnight rate from today through two years from today.

| Scenario	Probability	Average value of $_0i$
A (rise)	1/3	$5\frac{3}{4}$ (six months of 5 percent, eighteen months of 6 percent)
B (fall)	1/3	$4\frac{1}{4}$ (six months of 5 percent, eighteen months of 4 percent)
C (stay same)	1/3	5

$$_2i = \frac{1}{3} \bullet 5\tfrac{3}{4} + \frac{1}{3} \bullet 4\tfrac{1}{4} + \frac{1}{3} \bullet 5 = 5$$

$_3i$ is the expected value of the average value of the overnight rate from today through three years from today.

Scenario	Probability	Average value of $_0i$
A (rise)	1/3	$5\frac{5}{6}$ (six months of 5 percent, thirty months of 6 percent)
B (fall)	1/3	$4\frac{1}{6}$ (six months of 5 percent, thirty months of 4 percent)
C (stay same)	1/3	5

$$_2i = \frac{1}{3} \bullet 5\tfrac{5}{6} + \frac{1}{3} \bullet 4\tfrac{1}{6} + \frac{1}{3} \bullet 5 = 5$$
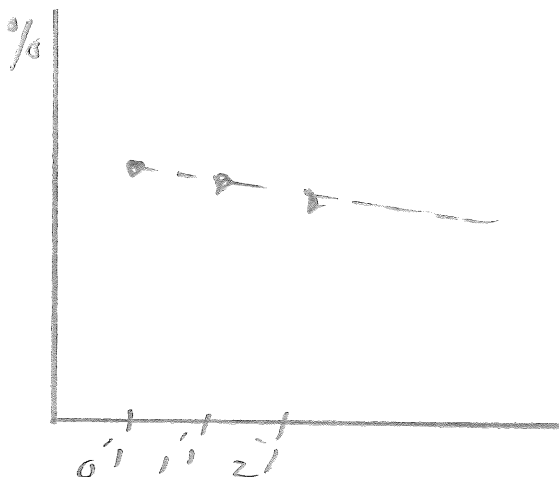
so the yield curve is flat:
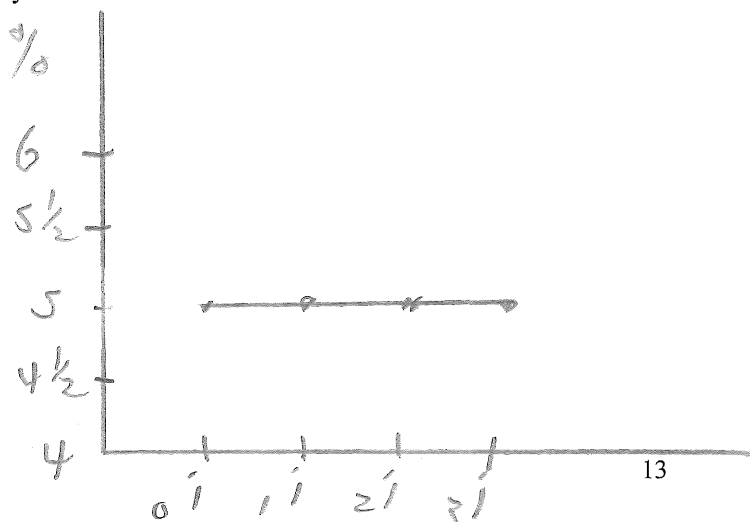
iv) Summary

Flat yield curve when market participants believe future overnight rates may be higher, lower, or the same as today's - no reason to expect any particular direction of movement. To put it another way, today's overnight rate is not especially high or low, relative to the likely future level of overnight rates.
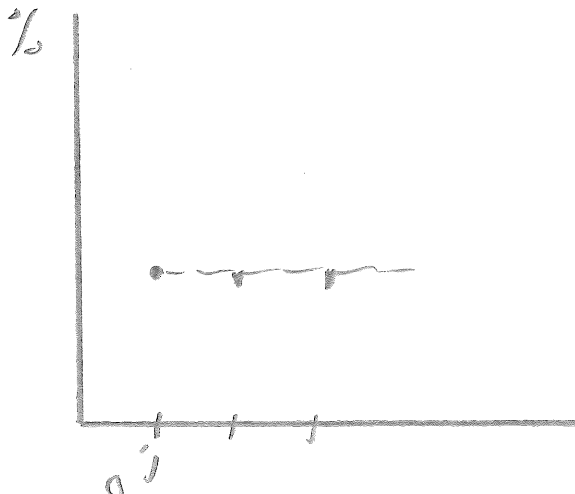
Upward-sloping yield curve when market participants believe future overnight rates may be higher than today's or the same as today's, but unlikely to be lower. To put it another way, today's overnight rate is especially low relative to the likely future level of overnight rates.
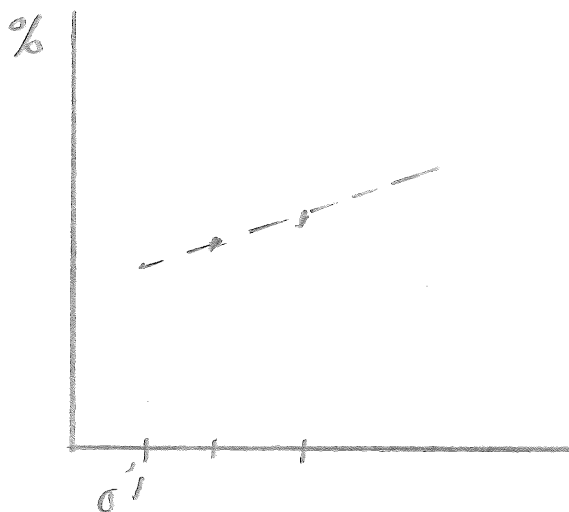
Downward-sloping yield curve when market participants believe future overnight rates may be lower than today's or the same as today's, but unlikely to be higher. To put it another way, today's overnight rate is especially high relative to the likely future level of overnight rates.

Notice that I drew the yield curves with dotted lines. In a minute, you'll see why.

D) Term premiums

1) A puzzle

I said at the beginning of this section that the expectations hypothesis is a useful part of our theory of the term structure, but is incomplete in itself because it is inconsistent with a well-know fact about the yield curve. Here is the fact that is inconsistent with the expectations hypothesis is: on *most* days the yield curve is upward-sloping; days when the yield curve is flat are fairly rare; days when the yield curve is downward sloping are quite rare. In fact a downward-sloping yield curve is so abnormal that is it called an "inverted" yield curve.

Why is this fact inconsistent with the expectations hypothesis? Recall that under the expectations hypothesis the yield curve is upward-sloping when people believe future overnight rates are likely to be higher than today's overnight rate - today's overnight rates are especially low relative to the likely future level of overnight rates. Thus, to explain the fact that the yield curve is usually upward-sloping using the expectations hypothesis, one would have to say that people *usually* think future overnight rates are likely to be higher than today's overnight rate. But that would be crazy. It is like saying that you usually think you will weigh more next week than you do today. That could only be true if you are getting fatter and fatter forever. It is like saying the usual weather forecast is that temperatures will be lower next month than they are today. That could only be true in a world that was getting colder and colder over time. (In fact, because of global warming the opposite is not so crazy, but you get my point.)

To solve the puzzle, we hypothesize that the yield on a bond of maturity m is equal to the expected value of the average value of future overnight rates over the next m years *plus* something. The extra something is called a "term premium."

2) Solution: term premiums

Term premiums can be defined like this. The yield on a one-year Treasury bond is equal to:

$$_1 i_t = E\left[\frac{_0 i_d + _0 i_{d+1} + _0 i_{d+2} + \ldots\ldots + _0 i_{d+365-1}}{365}\right] + {}_1\tau_t$$

The two-year yield is:

$$_2 i_t = E\left[\frac{_0 i_d + _0 i_{d+1} + _0 i_{d+2} + \ldots\ldots + _0 i_{d+729}}{730}\right] + {}_2\tau_t$$

The three-year yield is:

$$_3 i_t = E\left[\frac{_0 i_d + _0 i_{d+1} + _0 i_{d+2} + \ldots\ldots + _0 i_{d+1094}}{1095}\right] + {}_3\tau_t$$

Generally, the yield on a Treasury bond of maturity m is:

$$_m i_t = E\left[\frac{_0 i_d + _0 i_{d+1} + _0 i_{d+2} + \ldots\ldots + _0 i_{d+\ldots\ldots}}{m \bullet 365}\right] + {}_m\tau_t$$

$_1\tau_t, \, _2\tau_t, \, _3\tau_t, \ldots _m\tau_t$ are the term premiums. They mean that the yields are higher than the expected values of average future overnight rates. Notice that term premiums are different at different maturities. There is one term premium for one year bonds ($_1\tau_t$), a different term premium for two-year bonds ($_2\tau_t$), and so on.

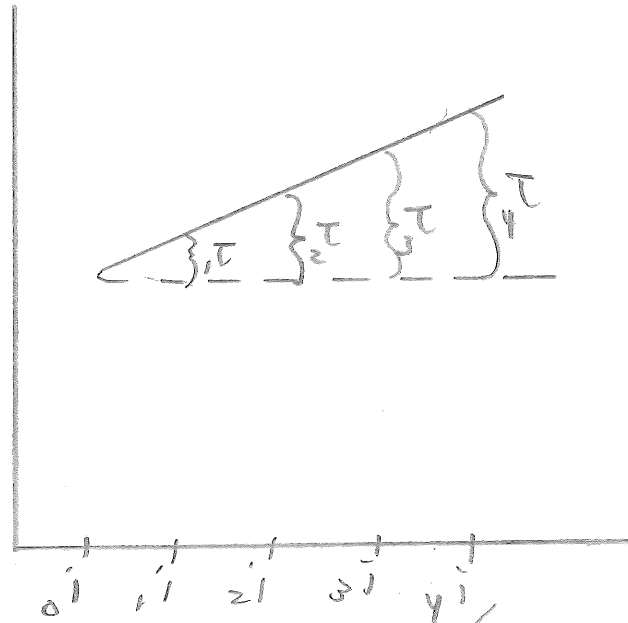Furthermore, term premiums tend to be *larger* for bonds of *longer* maturity. Thus:

$$_1\tau_t < \, _2\tau_t < \, _3\tau_t < \, _4\tau_t < \ldots\ldots$$

Finally, notice that our notation allows for the possibility that term premiums can change over time. $_1\tau_t$ can be different from $_1\tau_{t+1}$, and so on.
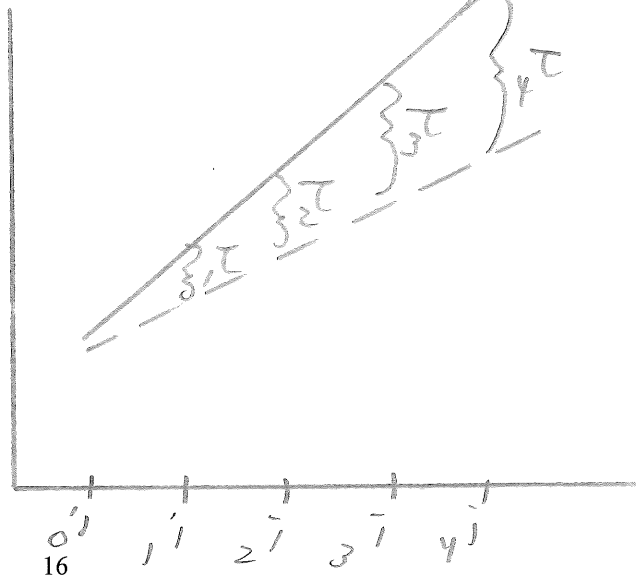
3) Yield curves with term premiums

How do term premiums solve the puzzle? Let me draw what yield curves will look like in each of the cases that I drew above, once I have added the term premiums to bond yields. As above, I will use dotted lines to draw the "expectations-hypothesis" yield curves, that is what the yield curves would look like *without* term premiums. I will use solid lines to draw what the yield curves look like *with* term premiums.
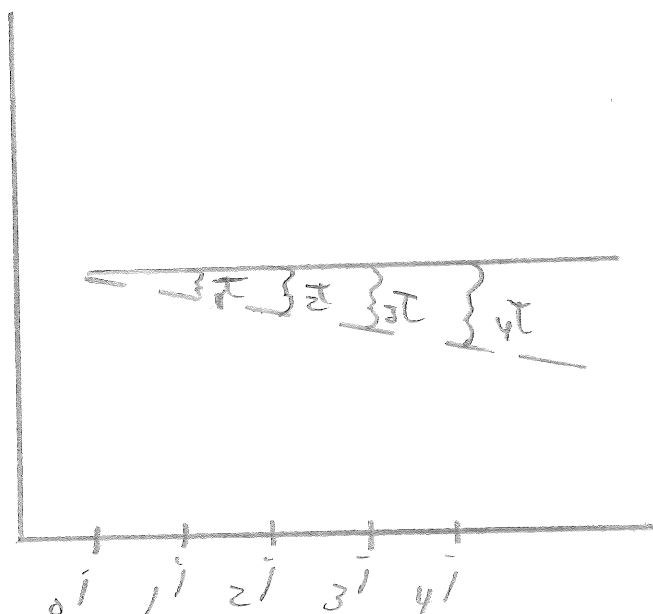
Market participants believe future overnight rates may be higher, lower, or the same as today's - no reason to expect any particular direction of movement. To put it another way, today's overnight rate is not especially high or low, relative to the likely future level of overnight rates.

Market participants believe future overnight rates may be higher than today's or the same as today's, but unlikely to be lower. To put it another way, today's overnight rate is especially low relative to the likely future level of overnight rates.

Market participants believe future overnight rates may be lower than today's or the same as today's, but unlikely to be higher. To put it another way, today's overnight rate is especially high relative to the likely future level of overnight rates.

Notice that once you put in term premiums, yield curves are upward-sloping *most of the time*, as they are in reality. The yield curve is upward-sloping *not only* when people think today's overnight rate is especially low relative to the likely future level of overnight rates, *but also* when today's overnight rate is not especially high or low relative to the likely future level of overnight rates.

4) Explanation of term premiums: interest-rate risk

So far, so good. But what explains the existence of term premiums?

Remember that generally a "premium" in the interest rate or yield on investment A relative to an alternative investment B is there because there is *something bad* about investment A relative to investment B. (Think about default-risk and liquidity - should be "illiquidity" - premiums).

What is bad about Treasury bonds relative to lending overnight? And what is especially bad about longer-term Treasury bonds relative to shorter-term Treasury bonds, so that term premiums increase with maturity?

Interest-rate risk! Any investor holding a bond, who thinks there is a chance she will have to sell the bond before it matures, must worry about interest-rate risk, that is the risk that an unexpected increase in bond yields will lower the price of the bond. Because of interest-rate risk, a bondholder can actually *lose money* holding a bond. And remember that interest-rate risk is greater for bonds of longer maturity. That is not true of overnight loans.

If the yield on a Treasury bond were just equal to the expected return to overnight lending, most investors would not buy Treasury bonds. They would choose to lend overnight instead. Prices of Treasury bonds must be low enough to give investors some extra yield, relative to overnight lending, as compensation for taking on interest-rate risk. That extra yield is term premiums.

©**Christopher Hanes Economics 450   Class notes**

VI) Liquidity
        A) Introduction and review
        B) What determines the bid-ask spread on a bond?
        C) Fire sale
        D) Why does it take more/less time to sell a type of bond at the highest possible price?
        E) General idea: the "lemons problem"
            1) Introduction
            2) Symmetric versus asymmetric information
            3) Knowledgeable buyers
            4) Lemons problem
        F) The lemons problem and liquidity of bonds
            1) Asymmetric information about default risk
            2) Lemons problems in bonds arise naturally
            3) Lemons problems in bonds are bad for everybody
            4) Things that can alleviate lemons problems in bond markets
                a) Bond ratings
                b) Financial disclosures
        G) Companies that cannot issue bonds

VI) Liquidity

A) Introduction and review

I hope you remember what I said about liquidity earlier. What I said was the following. "Liquidity" refers to the ease and speed with which you can convert an asset into cash. One component of liquidity is how long it takes to sell an asset for the highest possible price - I mean the highest price one is likely to get for that asset. Another component of liquidity is any work that must be done or monetary costs that must be paid to sell the asset. Relatively illiquid assets must pay higher returns. A "liquidity premium" is the extra interest rate or yield that an asset pays because it is relatively illiquid.

For loans, a monetary cost of selling the assets is fees to lawyers to arrange the transfer of a borrower's debt obligation from the original lender to the buyer of the loan.

For bonds, monetary costs of selling can be a fee paid to a bond broker to arrange the sale of a bond, or bond dealer's "bid-ask spread." A bond dealer, unlike a broker, buys and sells bonds herself. She makes a public offer to sell a bond at her "ask" price. The ask price is about equal to the highest price one is likely to get for that bond. She offers to buys the bond for a "bid" price. The bid price is lower than the ask price, so that the bond dealer can profit from the difference. Thus, a dealer's bid price must be lower than the price a bondholder would get from selling the bond directly to someone willing to pay the dealer's bid price. So why do bondholders sell bonds to dealers? Because it might take a bondholder a long time to find someone willing to buy at the higher price, and it might be trouble too. The bid-ask spread is in effect the cost to a bondholder of selling a bond *through* a dealer, that is selling to bond to the dealer who then sells it on to the ultimate buyer after taking a "cut." The dealer's cut is the bid-ask spread. Bid-ask spreads are bigger for some bonds than for others. The larger the bid-ask spread on a bond, the more costly it is to sell through a dealer, hence the less liquid is the bond.

All of that is what I said before. Now, in this section I return to the concept of liquidity to discuss a *fundamental determinant* of liquidity. The fundamental determinant of liquidity that I will discuss here affects *both* a dealer's bid-ask spread *and* the amount of time it would take for a bondholder to sell a bond himself. It is related to default risk, but it is not default risk *per se*. It is about the information about a bond's default risk that is available to potential bond buyers. It is specifically about the "symmetry" of information about default risk.

B) What determines the bid-ask spread on a bond?

Why are bid-ask spreads bigger for some bonds than for others? For a bond dealer, just as for a bondholder trying to sell a bond himself, some types of bond sell quickly. Other types of bonds are slow sellers. Bond dealers "charge" bigger bid-ask spreads on slow-selling bonds.

Why? First, understand that a bond dealer must hold an *inventory* of bonds to sell, like the inventory of goods held by a retailer. It is costly for a retailer to hold goods in inventory. Typically, a retailer borrows to get the money to buy his inventory of goods. (Retailers borrow from banks or by selling the kind of short-term bond called commercial paper.) The retailer pays off the borrowing with the proceeds of selling the goods. If the goods sell quickly, the retailer can pay off his borrowing faster, so he doesn't have to pay much interest on the loan. If the goods sell slowly, the retailer must borrow for longer, paying more total interest. Thus, holding an inventory of slow-selling goods is more costly for a retailer.

Similarly, bond dealers borrow to finance their inventories of bonds. Holding an inventory of slow-selling bonds is more costly for a dealer. The longer the average bond remains in inventory, the higher is the dealer's borrowing cost. Slow-selling bonds are more costly for a dealer to handle.

C) Fire sale

What happens if a bondholder or bond dealer is holding slow-selling bonds - bonds that take time to sell at the highest possible price - and the bondholder or bond dealer finds that she needs to get her hands on cash *immediately*? Then the bonds must be sold quickly, perhaps at very low prices.

In financial markets, such a situation is referred to as a "fire sale." This is when an illiquid asset must be sold quickly and therefore at a low price.

The phrase is an analogy to a fire sale in a retail business. This is when a retailer must sell off his stock of goods as fast as possible and therefore greatly marks down his prices. Such a situation can arise when a fire damages the building containing the store. Hence "fire sale."

D) Why does it take more/less time to sell a bond at the highest possible price?

For some bonds, there is not much variation in the price different people would be willing to pay for the bond. Just about everybody is willing to pay the same price. Those bonds will sell quickly at that price. A bondholder trying to sell the bond himself could sell it quickly. Bond dealers' bid-ask spread on the bond is small. These are relatively *liquid*. A

For other bonds, there is wide variation in the prices different people would be willing to pay for the bond. Some people would be willing to buy the bond at a relatively high price. But many other people would not pay that much for the bond. A bondholder trying to sell a bond himself must spend time and trouble to find one of the right buyers. A bond dealer selling the bond might have to wait a long time for one of the right buyers to come along, so dealers charge a large bid-ask spread on the bond. Those bonds are relatively *illiquid*.

What determines whether a bond falls more into the first category, or the second? This is where symmetry of information about default risk comes in.

E) General idea: the "lemons problem"

1) Introduction

There is a problem that can come up when people try to sell something - a good, a service, a financial asset such as a bond. It can make the thing illiquid. In extreme cases it can make it practically impossible to sell the thing at all. You might have heard about it in other economics classes. It is called the lemons problem. A lemons problem can occur when:
- units of the thing in question vary in quality: some units are good, some units are bad
- people *know* that the units of this thing vary in quality
- information people have about the quality of a particular unit is *asymmetric*
- everyone knows that information is asymmetric.

In this subsection I will lay out the general idea of a lemons problem. In the next subsection I will explain how it applies to bonds specifically.

2) Symmetric versus asymmetric information

Economists say that information is "symmetric" when everyone has the same information about something. Information is "asymmetric" when some people have relevant information that other people don't have. (This is an odd use of the words symmetric and asymmetric, I know, but it is the way economists use them here.)

Consider the thing that people are trying to sell. If everyone has the same information about the quality of an individual unit of the thing, then we would say that information about quality is symmetric. If information about quality is symmetric, then the thing will be easy to sell - it will be liquid. There will be many people willing to buy any given unit of the item. Everyone will be willing to pay about the same price for the unit. The thing will be liquid.

One way that information could be symmetric is if *everyone* can easily observe the quality of an individual unit of the item. In this case, a buyer and seller will agree to a low price if the item is bad, a high price if the item is good. Both good and bad items can be sold quickly.

Another way that information can be symmetric is if *no one* can observe the quality of the item to be sold. You might guess that this situation is bad, but actually it's fine - the market will still function well. Neither the buyer nor the seller knows the quality of the good. Because they have the same information, they place the same probability on the item being bad. They will quickly agree to a price that reflects that probability. A buyer will be happy if the item turns out to be good, sad if the item turns out to be bad. But again all items can be sold quickly.

What if some people have better information about the quality of an individual unit than other people have? Then information about quality is asymmetric.

3) Knowledgeable buyers

When information about quality of a thing is asymmetric, some people will have relatively good information about the quality of an individual unit. A person who has relatively good information, and *knows* that she has relatively good information, is called a "knowledgeable buyer."

4) Lemons problem

The "lemons problem" comes up if information about the quality of the item is asymmetric, and people with relatively bad information know that there are other people who have better information.

Under these circumstances a person who doesn't know the quality of the item will suspect that anyone trying to sell a unit is trying to sell that unit because the seller knows the unit is a bad one - if it is a good unit, why wouldn't the seller want to keep it? So the people who don't know the quality of the item will not buy any units, or buy only at very low prices.

Anyone trying to sell a high-quality unit of the item will want to find a knowledgeable buyer. A knowledgable buyer is not afraid of being rooked by the seller, and will be willing to pay a relatively high price for a good unit of the item. If there aren't many knowledgable buyers, it could take a long time to sell the item at the price a knowledgable buyer would be willing to pay. The thing will be illiquid. (If there are almost no knowledgeable buyers then the thing might be practically unsellable except at a low price.)

A market like this is called a "market for lemons."

What happens if a seller in a market for lemons needs to sell quickly? Too bad for him! He may not find a knowledgable buyer in time. Then he will have to sell the item at a very low price.

F) The lemons problem and liquidity of bonds

1) Asymmetric information about default risk

In the case of a bond, "quality" means default risk - the risk that the issuer of the bond will not pay the IOU(s) as promised. For some bonds, information about default risk is symmetric - this can be because no one has much information to assess default risk, or because everyone has lots of information to assess default risk. For these bonds, there is no lemons problem. Lots of people are willing to buy the bond at about the same price. The bond is liquid.

A lemons problem arises in sales of a bond when there is asymmetric information about the bond's default risk. The lemons problem makes the bond a slow seller, hence illiquid. A bondholder trying to sell the bond himself must spend time and trouble to find a knowledgable buyer. A bond dealer charges a large bid-ask spread on the bond. Because the bond is illiquid, its yield is relatively high. Its yield contains a liquidity premium.

Note that default risk *per se* does not cause the lemons problem. A bond can have high default risk and still be very liquid, as long as:
- many people have the same information about the bond's default risk
- they know that other people have the same information about the bond's default risk.

A lemons problem arises if potential bond buyers suspect that there are some people in the market who have special, extra information about the bond's default risk.

2) Lemons problems in bonds arise naturally

It is easy for a lemons problem to arise in sales of a bonds issued by companies. A company will have to default on its bonds if it does not generate enough revenue to pay the IOUs, and goes bankrupt. Some people, such as a company's managers, their family and friends, are naturally going to be the first to know if a business is beginning to fail, or is in a dangerous situation. That means information is asymmetric.

3) Lemons problems in bonds are bad for everybody

Other things equal, yields are higher on less-liquid bonds (liquidity premiums). Thus, a firm can benefit from avoiding lemons problems in its bond issues: that would allow the firm to borrow more cheaply by lowering the yield on its bonds.

For society as a whole, reducing liquidity premiums generally lowers the cost of borrowing for firms giving firms incentive to install more capital equipment, boost labor productivity and hence raise real wages in the economy.

4) Things that can alleviate lemons problems in bond markets

a) Bond ratings

Remember bond ratings: they are "grades" given to bonds by companies such as Moody's and Standard and Poor's, which indicate the bond rater's estimate of the bond's default risk. I told you earlier that a company issuing a bond will pay bond raters to rate the bond and publish the rating, even if the rating is low. Now you know why. A bond rating makes information about default risk on the bond less asymmetric. Some people who would not buy the bond without a rating will buy it with a rating. Thus, the bond is more liquid; its liquidity premium is smaller; its yield is lower. The company can borrow more cheaply.

b) Financial disclosures

A company is called "publicly owned" if it has sold shares of stock. In the United States, publicly owned companies are required to publish a great deal of information about the state of the company's business. There are severe penalties for lying or hiding information. The system is overseen by a federal government agency, the Securities and Exchange Commission or "SEC."

4

These mandatory financial disclosures make information about publicly held companies' default risk less asymmetric, and hence lowers yields on bonds issued by publicly held companies.

G) Companies that cannot issue bonds

Many companies cannot issue bonds because they can't get around the asymmetric information problem. People would not buy a bond issued by that company, or would buy it only at very low price relative to the IOU..

These companies include many privately held companies, which are not subject to the SEC's financial disclosures system, and newly established companies for which bond raters cannot estimate default risk.

Companies that cannot issue bonds must borrow through loans from banks and other lenders, even though interest rates on loans are higher than most bond yields.

©Christopher Hanes   Economics 450   Class notes

VII) Loans
        A) Introduction
        B) Bankruptcy
                1) What it is
                2) Bankruptcy limits the downside to bankrupt people or businesses
        C) General problem: borrowers have incentive to do high-risk, high-return things
                1) Problem: moral hazard
                2) Economists' jargon: moral hazard
                3) Solutions
                        a) Lend only to borrowers with something to lose
                        b) Collateral
        D) More specific problems
                1) Introduction
                2) Before the loan is made: asymmetric information about potential borrowers' projects
                        a) The problem
                        b) Something that won't work: just charge a higher interest rate
                        c) Solution: screening
                3) After the loan is made: borrower's behavior
                        a) A borrower can carry out approved project in risky way
                        b) Solution: monitoring, restrictive convenants
        E) Why loans are relatively illiquid assets
                1) Introduction
                2) Screening and monitoring create asymmetric information
                3) Power to enforce restrictive covenants must be transferred
        F) Summary
                1) Why interest rates on loans are relatively high
                2) What determines interest rates and yields

VII) Loans

A) Introduction

At the end of the previous section, I said that some firms can't issue bonds, and must therefore borrow through loans, because there is asymmetric information about the risk that the firm would default on a bond it issued. But a firm that defaults on bonds would default on loans, too. Why wouldn't asymmetric information about default risk prevent a firm from borrowing through loans?

In this section, I will tell you about the business of lending. I will explain some techniques lenders use to mitigate the risk that a borrower will default on a loan.

B) Bankruptcy

1) What it is

Bankruptcy is a legal process in which a firm or person declares itself/himself unable to pay all of its/his debts. A court seizes the bankrupt's assets. The court liquidates the seized assets and uses the money to pay off the bankrupt's debts as much as possible. There won't be enough money to pay everyone, of course - that is why the bankruptcy happened. Laws and legal precedents determine which creditors get first dibs on the money that is available.

Once the court has allocated the available money to the creditors, the bankrupt's remaining debts are cancelled. A person who went bankrupt, or the people who controlled a firm that went bankrupt, cannot be required to pay those unpaid debts in later years.

2) Bankruptcy limits the downside to bankrupt people or businesses

Notice that bankruptcy limits the loss to a bankrupt person or owners of a bankrupt firm. Their assets are brought down to zero, but then they are free to start other businesses. Zero is the bottom.

C) General problem: borrowers have incentive to do high-risk, high-return things

1) Problem: moral hazard

The fact that bankruptcy limits the downside to a borrower gives borrower incentive to do high-risk, high return things.

As an example, suppose I have borrowed $500,000, and I get an opportunity to do something - a business strategy I can follow, or a specific action I can take - which has a high probability of destroying my business, putting me into bankruptcy, but a low probability of paying off BIG - like, $10,000,000. This opportunity is, in effect, a gamble I can take. The gamble consists of a large probability my wealth will be brought down to zero, along with a small probability I will become stinking rich. I might well choose to take the gamble!

Note that the person who lent me the $500,000 definitely does *not* want me to take this gamble. If it goes badly, the lender loses the $500,000 plus interest. If it goes well and I get the $10,000,000, the lender gets none of that big upside; the lender just gets the $500,000 plus interest.

Economists use the term "moral hazard" to describe a situation where one party to a contract or transaction has incentive to take risks that can injure the other party. (The term can mean other things, too.) This situation is an example of moral hazard.

Things would be different if the penalty for going bankrupt were greater - for example, if a bankrupt person or manager of a bankrupt firm were put into prison until all his debts are paid off. I mention this prison thing because it's what we actually did to bankrupts at times in history. But we got rid of "debtors' prisons" a long time ago.

2) Solutions

a) Lend only to borrowers with something to lose

To mitigate this moral hazard problem, lenders are careful to lend only to a person or company with sufficient "net worth." Net worth means the value of assets minus existing liabilities (debts). A potential borrower with positive net worth has something to lose in a bankruptcy.

Notice that this means a person or business with little net worth may not be able to borrow, even if he/it has potentially profitable low-risk business opportunities. It also means that a person or business may lose the ability to borrow if the value of his/its assets falls, reducing his net worth.

b) Collateral

Frequently, a loan contract will specify one or more of the borrower's assets as collateral for the loan, which can be seized by the lender if the borrower fails to pay as promised. That is useful to a lender because in some cases the lender can seize the collateral without all the rigmarole of bankruptcy. Even if the borrower goes into bankruptcy, having something specified as collateral in the loan contract may mean the lender will not have to stand in line with the rest of the creditors hoping to get paid something out of the bankrupt's asset liquidation.

D) More specific problems

1) Introduction

The general problem - bankruptcy can give borrowers incentive to do high-risk, high return things - takes two specific forms. These more specific problems have additional specific solutions.

2) Before the loan is made: asymmetric information about potential borrowers' projects

a) The problem

Some people who want to borrow from a lender have low-risk projects in mind. Some of these low-risk projects may have high returns. Some may have low returns - just enough to pay off the loan with a little profit left over for the borrower. Either way, the lender gets paid. Thus, it is good for a lender to make a loan to a borrower with a low-risk project.

But some people who want to borrow from a lender have high-risk projects in mind. If the unlikely, but possible upside on a high-risk project is very profitable, it may be a good gamble for the person with the high-risk project, even though the downside is bankruptcy. But the lender doesn't want to lend to people with such projects.

A borrower knows what kind of project he has in mind. But the lender does not. This is an example of asymmetric information.

If the lender posts an interest rate and offers to lend to anyone who wants to borrow at that rate, many of the borrowers may be people with high-risk, high-upside projects, and many of these will default. The lender may end up losing money on the lending operation.

What can a lender do?

b) Something that won't work: just charge a higher interest rate

One might think the solution to this problem is for the lender to post a high interest rate, high enough to make lending profitable even if there are lots of defaults. But that won't work. Posting a high interest rate will actually raise the fraction of loans that end up in default. That is because the potential borrowers with low-risk, low-return projects won't want to borrow at a high interest rate. Their projects will throw off some profit for the borrower if theh loan is at a low interest rate, but not if the loan is at a high interest rate. Thus, if the lener posts a high interest rate, the people who still want to borrow from the lender will mostly be the one with high-risk, high-return projects who are likely to default on their loans.

c) Solution: screening

To solve the problem, the lender needs to make information more symmetric. The lender needs to collect information about the people trying to borrow and their projects, enough to identify most if not all of the high-risk projects, and *not* lend to them. This is called "screening."

Of course, screening is costly for the lender. The lender must hire smart people to collect the information and analyze it. But this is a necessary cost of running a lending business. It must tend to increase the interest rate charged on loans, because the interest rate must cover the costs of screening to make the lender willing to make loans rather than just buy bonds.

3) After the loan is made: borrower's behavior

a) A borrower can choose to carry out an approved project in risky way

After a loan is made, a borrower who got through the screening process may still choose to carry out his project in a high-risk, high-upside way. This is, again, an asymmetric information problem.

b) Solution: monitoring and restrictive convenants

The solution here has two parts.

One part is to make the information more symmetric by continuing to collect information about the borrower after the loan is made. This is called "monitoring." But monitoring is not enough by itself. The borrower already has the money. So the lender catches him doing something risky. So what?

What the lender must do is add terms to the loan contract so that the borrower is penalized if the lender observes him doing something risky. These contract terms are called "restrictive convenants." Restrictive covenants typically list some things the borrower may not do and the resulting penalty.

Monitoring and restrictive covenants, like screening, involve costs for the lender. People must be hired to do the monitoring. Lawyers must be paid to enforce the contract terms if the borrower tries to violate them. But again, this is a necessary cost of running a lending business. Like costs of screening, costs of monitoring and enforcement of restrictive covenants must tend to increase the interest rate charged on loans, because the interest rate must cover these costs (as well as costs of screening) to make the lender willing to make loans rather than just buy bonds.

E) Why loans are relatively illiquid assets

1) Introduction

I have often said that loans are less liquid than bonds. What that means is, it can take a long time for an original lender to find someone willing to buy a loan, and there can be high costs of transferring the borrower's debt obligation from the original lender to the person buying the loan. Now that you have learned more about what loans involve, it's easy to see some reasons for that.

2) Screening and monitoring create asymmetric information

When the original lender screened the borrower before making the loan, the original lender collected hard-to-get information about the probability the borrower would default on the loan. As the original lender monitored the borrower's behavior after the loan was made, the original lender collected still more information about the borrower. All of this information is hard for a potential buyer of the loan to get. Thus, potential buyers of a loan will often have reason to fear that the original lender has better information about the risk of default on the loan. Asymmetric information!

So there can be a lemons problem when it comes to selling a loan. A possible buyer may suspect the original lender wants to sell the loan because the oroginal lemnder knows the default risk is high. That can make it hard for the original lender to find someone willing to buy the loan, except at a low price. The original lender must find

3) Power to enforce restrictive covenants must be transferred

I said that a loan may be relatively illiquid partly because of the monetary costs from fees to lawyers arranging the transfer of the borrower's contractual debt obligation from the original lender to the buyer of the loan. It is also necessary to transfer the original lender's power to enforce restrictive covenants - more legal work.

F) Summary

1) Interest rates on loans are relatively high

One expects interest rates on loans to generally be higher than yields on bonds of matching maturity, even if the true default risk on the loan is the same as the default risk on the bond, because:

- loans are relatively illiquid. The interest rate on a loan contains an especially big (il)liquidity premium.

- the interest rate on a loan must cover the costs of screening, monitoring and enforcing restrictive covenants.

2) What determines interest rates and yields

All interest rates and yields in an economy are determined by:

- current and expected future overnight rates
- term premiums
- default-risk premiums (except for Treasury bonds and perhaps a few other types of bonds)
- liquidity premiums (except for Treasury bonds and perhaps a few other types of bonds)
- tax treatment (for municipal bonds)

In terms of the notation we have been using, the yield on a Treasury bond of maturity m is:

$$_m i_t = E\left[\frac{_0 i_d + {}_0 i_{d+1} + {}_0 i_{d+2} + \ldots\ldots + {}_0 i_{d+\ldots\ldots}}{m \bullet 365}\right] + {}_m \tau_t$$

The yield on a zero-coupon corporate bond of maturity m is:

$$_m i_t = E\left[\frac{_0 i_d + {}_0 i_{d+1} + {}_0 i_{d+2} + \ldots\ldots + {}_0 i_{d+\ldots\ldots}}{m \bullet 365}\right] + {}_m \tau_t + default\,risk\,premium + liquidity\,premium$$

The interest rate on a comparable loan may be higher still to cover costs of screening etc.

©**Christopher Hanes   Economics 450   Class notes**

VIII) Money, liquidity preference and determination of interest rates in a simple economy
    A) Introduction
    B) Money
        1) Introduction
        2) Trades without money (barter) and double coincident of wants
        3) Medium of exchange
    C) Types of money
        1) Introduction
        2) Commodity money
        3) Fiat money
            a) Definition
            b) Types of fiat money
                i) Paper money
                ii) Token coins
                iii) Ledger money
    D) Holding money
        1) Definition
        2) Opportunity cost of holding money
        3) Interest paid on money
        4) Why choose to hold money? Liquidity preference
    E) The money demand function
        1) The money demand function
        2) The money demand function
            a) The variables that determine the quantity of money demanded
            b) The general relationship between the variables and M^D
    F) Money demand, money supply and the market-clearing interest rate i
    G) The Baumol-Tobin Model
        1) Introduction
        2) Story and notation
        3) Graphs to describe M_d and M, B_d and B.
        4) Possible actions the person could take
            a) Hold just bonds, no money
            b) Hold just money, no bonds
            c) On the first day, put half his income in bonds, half in money
            d) On the first day, put two-thirds of his income in bonds, one-third in money
            e) A generalization
        5) How to find the answer with calculus
            a) The method
            b) Applying the method here
            c) Checking the answer
        6) Another way to set up the problem: minimizing cost
   H) How is our economy different from the one we have examined in this section?

VIII) Money, liquidity preference and determination of interest rates in a simple economy

A) Introduction

We have gone a long way toward explaining the entire structure of interest rates in an economy, taking "interest rates" to mean bond yields as well as interest rates on loans. Given current and expected overnight rates, we can determine all the rest by adding in term premiums, liquidity premiums and default-risk premiums.

But what determines overnight rates? In this section, I will present a theory of the way the overnight rate would be determined in a simple economy. By "simple economy," I mean an economy where there are loans, bonds and money, but no banks and no central bank.

The theory I present should be familar to you.
It is presented in most
intermediate macroeconomics classes,
such as Econ 362 here at Binghamton.
It is about supply and demand for
"real money balances." I mean this thing:
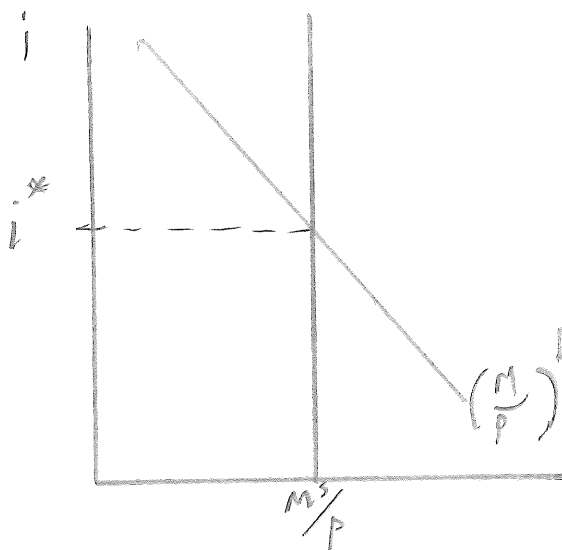
But look out! I will add something to what you learned in Econ 362. What I will add is a specific mathematical model of demand for real money balances. The mathematical model is called the "Baumol-Tobin" model. It will require you to use calculus and do some algebra.

Also, do not forget that this theory (including the Baumol-Tobin model) does *not* describe an economy like the one we live in. Our economy has banks and a central bank. In a later section you will learn how the overnight rate is determined in our economy. Yes, I know that in Econ 362 you were told that the money supply-money demand model *does* apply to our economy. That was a lie. Sorry.

B) Money

1) Introduction

I begin by defining the term "money" as used by economists. Money is a social construct that makes it easier to trade goods and services. To understand it, you first have to think about how people would trade things if money did not exist.

2) Trades without money (barter) and double conicidence of wants

"Barter" is when people directly exchange one good (or service) for another. In order for a barter to take place, person A has to have something person B wants, *and* person B has to have something person A wants. This condition is called "double coincident of wants."

Because the double coincidence of wants condition must be satisfied, it can take a lot of time to carry out trades using barter.

As an example, suppose you have eggs and want to get milk. You have to find someone who has milk and wants eggs. Perhaps you can't find anyone like that. But do you find a person who has milk. So you ask him: "what will you take in exchange for your milk?" He might say "beer." You might then go around trying to find someone who has beer and wants eggs (which is what you have, remember). Perhaps you can't fund anyone like that, but you do find a person who has beer. So you ask her.....You see why this can take time, and effort.

2

3) Medium of exchange

A medium of exchange is a thing that a society uses as an in-between good for all trades. Suppose, for example, that the medium of exchange is cupcakes. Then if you have eggs and want to get milk, you take cupcakes in exchange for your eggs, and anyone who has milk will take cupcakes in exchange for it.

The thing a society uses as its medium of exchange - in my example, cupcakes - is its money. Money may serve other purposes as well (e.g. "store of value," "unit of account" - you don't need to worry about those). But it is the medium of exchange function that makes a thing money.

When a society has a medium of exchange, that is when it uses money, trades can take place without double coincidence of wants.

C) Types of money

1) Introduction

All kinds of things have been used as money at one time or another. Two general categories of things used as money are "commodity money" and "fiat money."

2) Commodity money

This is what we call it when a society chooses some useful thing, a thing that would have value if it weren't money, and *also* uses it as money.

Here are examples of things that have been used as commodity money: cows; cigarettes; nails; copper; iron; gold; silver.

When people uses a metal such as copper, gold or silver as money, they may pass around lumps or ingots of the metal and define prices by weight, e.g., "how much do you want for that cow?" "six ounces of silver." But constantly weighing metal is a chore. So most societies that use metal money have stamped lumps of it into *coins* of standard weight, guaranteed by some trusted institution. For example, a coin might contain one ounce of silver, so to pay six ounces for the cow, you need only count out six coins. But the value of the coin still comes from the weight of metal that is in the coin. A coin that is visibly worn and under weight will buy less stuff than a full-weight coin.

3) Fiat money

a) Definition

This is what we call it when a society takes a useless or almost-useless cheaply produced thing as money - something that would have no, or very little, value in trade if it weren't being used as money.

b) Types of fiat money

i) Paper money

Here's an example: greenish pieces of paper printed with pictures of presidents.

ii) Token coins

This is when a society makes coins out of a very cheap metal like zinc. As money, the coins trade for values much higher than the cost of the cheap metal used to make them. It does not matter if a token coin is underweight

iii) Ledger money

Money does not have to be a physical object. Rather than pass around pieces of paper or token coins, a society keep a spreadsheet in a ledger - an account book or a computer program such as Excel. The spreadsheet lists everyone by name and, beside each name, a number that is the number of money units owned by that person. When person A sells something to person B, the agreed-upon number of money units, representing the agreed-upon price, is subtracted from person B's balance, added to person A's balance.

In the United States today, some money is paper money and token coins, but most money is ledger money. The ledger is kept on a computer system managed by the Federal Reserve. This ledger money is held by banks, not people. I'll get back to this later.

D) Holding money

1) Definition

Economists mean something specific when they say a person is "holding money." Holding money is distinct from *using* money. A person could *use* money to buy and sell things without holding money. How could you use money without holding money? Like this: whenever you sell something and get some money, you *immediately* use that money to buy bonds or make loans. You don't keep any money around for longer than it takes to do that. When you plan to buy something, you sell bonds or call in some overnight loans *the moment before* you buy the thing.

Holding money means hanging on to some money at times when you have no immediate plans to buy something, and haven't sold anything for a while - keeping money in your pocket or in a box for substantial spans of time in between buying and selling things.

2) Opportunity cost of holding money

I hope you remember the term "opportunity cost." Opportunity cost is what you give up when you choose to do something - your immediate alternative choice.

There is an opportunity cost of holding money. It is the interest you could earn on the most liquid interest-paying asset available to you. If you can make overnight loans, it might be the interest rate on overnight loans. Or it could be the rate of return to buying and selling liquid bonds. Why is this the opportunity cost of holding money? If you use money but don't hold money, in the way I just explained, you can hold more bonds or make more loans. The extra assets that you would choose to buy if you didn't hold money would probably be the most liquid ones available to you, because you would need to buy and sell them almost every day to get money to buy stuff. Thus, by choosing to hold money for substantial periods of time, you are giving up the opportunity to hold more liquid bonds or make overnight loans. You are giving up the return on the liquid investments you aren't making.

3) Interest paid on money

For most of history, money paid exactly zero interest. This was true for nearly all forms of commodity money. Think of gold, for example. If the money in your society is gold, and you choose to hold 9 ounces of gold for a month, at the end of the month you will have exactly 9 ounces of gold - not 9 ounces plus x percent of 9 ounces. You will have received zero interest on your gold.

Things are different nowadays. Paper money pays no interest, but the ledger money that is kept on a Federal Reserve computer system *does* pay interest (most of the time). We'll get back to that later.

But remember, in this section of the course, we are talking about a simple economy with no banks and no central bank. In economies like that, money pays no interest. So for this section, we'll assume money pays zero interest.

4) Why choose to hold money? Liquidity preference

Since money pays no interest in the economy we're talking about, and a person holding money incurs the opportunity cost of lost interest, why would anyone choose to hold money? Why not use money without holding it by frequently buying and selling bonds or making and calling in overnight loans?

Economists have a theory to explain that. It is called "liquidity preference" theory. The theory says that people hold money because two conditions hold. First, you need money to buy and sell things.You cannot trade interest-paying assets directly for goods and services. You must "liquidate" the interest-paying asset, that is turn it into money, e.g. by selling bonds, then use the money to buy goods and services. Second, even the most liquid interest-paying assets are at least a *little bit* illiquid in some way - it takes at least some time and/or it costs something to turn the asset into cash (e.g. the dealer's bid-ask spread on bonds). As long as these two conditions hold, money will be held even if the most liquid assets pay a higher return than money does.

E) The money demand function

1) The money demand function

Under the conditions assumed by liquidity preference theory, not only will people choose to hold money, but the *amount* of money that people choose to hold, that is the *quantity demanded* of money, depends on certain variables in a certain way. The relationship between these variables and the quantity of money demanded is called the "money demand function."

2) The money demand function

a) The variables that determine the quantity of money demanded

Let M^D denote the number of money units people want to hold - the quantity demanded of money. If the economy uses paper fiat money, it is the number of dollars or whatever that people want to hold. If the economy uses gold as money, M^D would be the number of ounces of gold people want to hold for monetary purposes (not for jewelry or whatever).

Liquidity preference theory implies that M^D depends on at least three things:
- P the price level. M^D is positively related to P: an increase (decrease) in P tends to increase (decrease) M^D, other things equal.
- Y real income or real spending. M^D is positively related to Y: an increase (decrease) in Y tends to increase (decrease) M^D, other things equal
- i the overnight interest rate or rate of return on the most liquid interest-paying asset or assets i. (But remember, this asset is not *perfectly* liquid.) M^D is negatively related to i: an increase (decrease) in i tends to decrease (increase) M^D, other things equal.

Also, there is something special about the relationship between the price level and money demand: changes in the price level affect the quantity demanded of money *proportionally*. That is, if the price level goes up (down) ten percent, the quantity demanded of money goes up (down) by exactly ten percent. This is because people hold money to buy stuff. If you need ten percent more money units to buy the same amount of real stuff, people will want to hold ten percent more money units.

VIII

b) The general relationship between the variables and M^D

The money demand function in its most general form is:

$$M^D = P \bullet L(i, Y)$$

This expression says that M^D is equal to the price level times a mathematical function called "L." At this point we don't want to be specific about what L looks like, but we do want to say that two variables enter into L: i and Y. But we do want to say that M^D is negatively related to i, positively related to Y. To say that, we can write:

$$M^D = P \bullet L(\overset{-}{i}, \overset{+}{Y})$$

Notice that the expression implies that changes in the price level affect the quantity demanded of money exactly proportionally. Now divide both sides of the equation by the price level P. That gives:

$$(M^D / P) = L(\overset{-}{i}, \overset{+}{Y})$$

The ratio of M to P (M/P) is called "real money balance" (or "real money balances"). By the way, why did we call the function of i and Y "L"? That is traditional. L is for "liquidity preference."

But why does the money demand function takes this form? And what is the specific mathematical function for L? To answer those questions, one needs a model. There are many models that illustrate the general notion of money-demand liquidity preference. All the models are consistent with the general function above, but the specific mathematical function varies from model to model. Also, depending on the model, there may be additional variables that affect money demand. Later on, I will show you one of those models, the "Baumol-Tobin" model. But before I do that, let's finish up the story from Econ 362 and show how the interest rate i is determined.

F) Money demand, money supply and the market-clearing interest rate i

I hope you remember this part from Econ 362, because I'm not going to spend much time on it. You plot out the money-demand function, in terms of real money balances, on a graph. You put the supply of real money balance on the same graph. The overnight rate in the economy (that is the rate on overnight loans, or the overnight return to holding liquid bonds) is the rate that lets the demand for real money balances equal the supply.
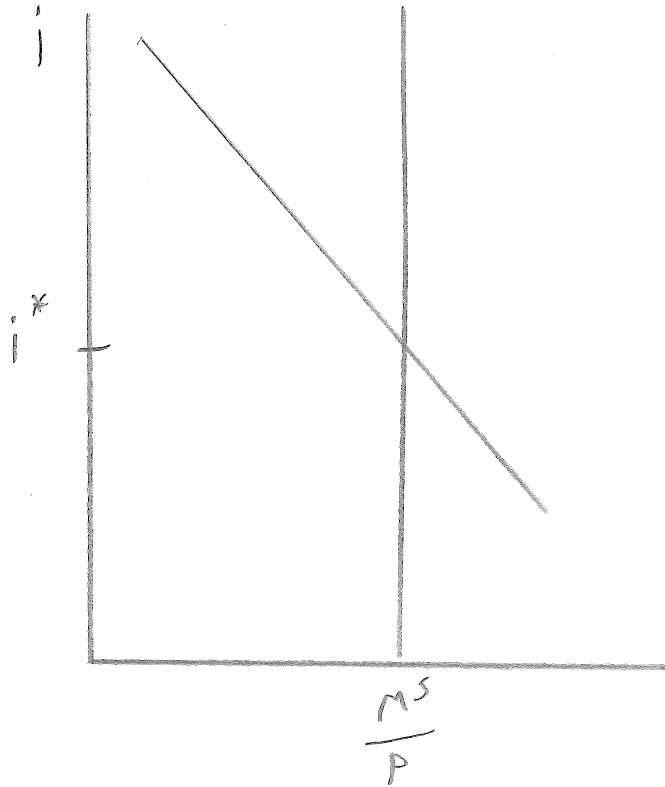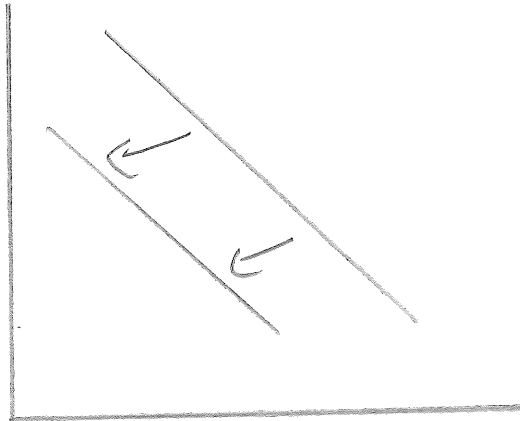
The money-demand function is shifted by changes in Y. An increase (decrease) in Y shifts the demand function out (in).
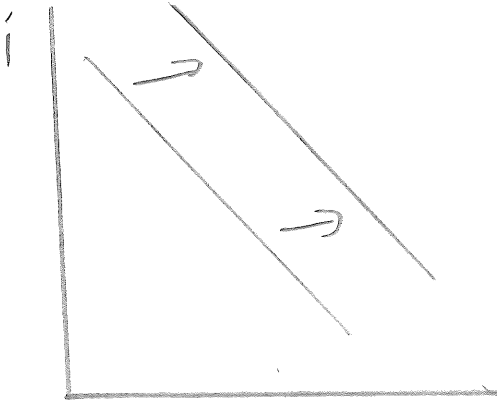
Increase in Y   *Decrease in Y*

Important: changes in the price level P do *not* shift the money-demand function, because we have defined it in terms of demand for *real* money balances (M^D / P). (If the number on the horizontal axis were M rather than M/P, changes in the price level *would* shift the demand curve.)

Meanwhile the supply of real money balances is determined by the price level P and the supply of money units M^S. If the economy uses gold as money, M^S would be the number of ounces of gold in the economy available for use as money (not for jewelry or whatever). If the economy uses paper fiat money, it is the number of paper dollars in the economy. The money-supply line is shifted by changes in P or M^S.

An increase in M^S holding P fixed shifts $\dfrac{M^S}{P}$ out:
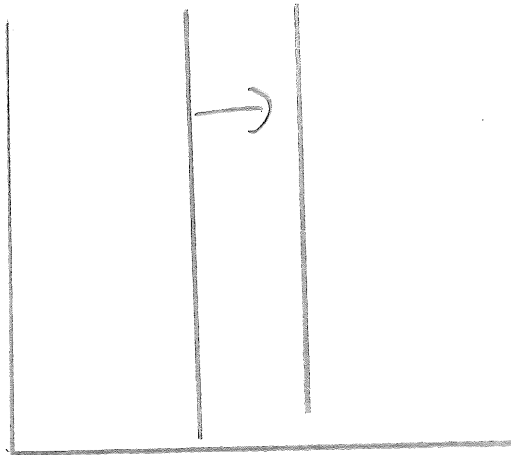
An increase in P holding M^S fixed shifts $\dfrac{M^S}{P}$ back:
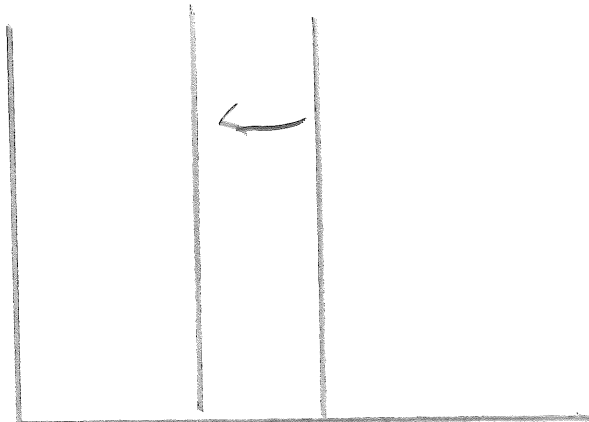
G) The Baumol-Tobin Model

1) Introduction

The Baumol-Tobin model is the simplest model of money demand due to liquidity preference. It is called the Baumol-Tobin model because it was invented by two people at about the same time, (William) Baumol and (James) Tobin.
The model illustrates one specific way that the key conditions of the liquidity preference theory can imply a money-demand function consistent with $(M^D / P) = L(i, Y)$. Remember those key conditions are:
- trades are made using money
- all assets interest-paying assets, even the most liquid ones, are at least a little bit illiquid.

In the Baumol-Tobin model, what makes the assets illiquid is a cost of turning them into cash, sort of like the bid-ask spread on selling bonds. Remember that another way an asset can be illiquid is that it takes *time* to cash out of the asset. There are models of liquidity preference in which assets are illiquid in that way. Those models also give money-demand functions consistent with $(M^D / P) = L(i, Y)$. But they are more complicated than the Baumol-Tobin model.

The model describes one person's demand for real money balance. To get the total money demand in the economy, you would say that everyone is like the person described by the model.

We are not interested here in the effect of changes in the price level, so to keep the notation simpler we will say that the price level P is unchanging and equal to one. That means the quantity of money M is the same as the quantity of real money balance M/P.

We want to derive a *mathematical function* that describes the relationship between the amount of M that the person chooses to hold and the values of the variables that affect his choice, especially i and Y. That function is the "money demand function." We want to see if the function is consistent with the general statement:

$$(M^D / P) = L(\overset{-}{i}, \overset{+}{Y})$$

We also want to see if there are any variables that affect M in addition to i and Y.

2) Story and notation

The person in the model receives a payment of real income at the beginning of a regular calendar period, say a year. The amount of the person's real income is Y. The person buys the same things, at the same price, every day, using up his entire annual real income Y over the course of the year. Thus, Y is the person's annual real spending as well as his annual income. On each day, the person spends $(Y / 365)$.

The person can hold money, which pays no interest. You will see that the person will choose to hold a different amount of money from day to day. We will use M_d to denote the quantity of money (number of money units, e.g. number of dollars) the person holds on day d. M_1 is the money he holds on the first day of the year, M_{365} is the money he holds on the last day of the year, and so on. We will use M (no subscript) to denote the *average* quantity of money he holds across all days of the year. This is what we are interested in. It is what we call the person's quantity demanded of money.

The person can also hold an interest-paying asset. This is the most liquid asset available. The asset pays a daily return that adds up to i over a whole year. That is to say, if you hold the asset for just one day, the interest you get is $\dfrac{i}{365}$. You can think of $\dfrac{i}{365}$ as being the daily interest paid on overnight loans, or the rate of return on buying bonds and selling them

after one day. I will call the asset "bonds." You will see that the person will hold a different amount of bonds from day to day. We will use B_d to denote the person's holdings of bonds on day d. B_1 is the bonds he holds on the first day of the year, and so on. B is the *average* daily investment in bonds across all days of the year.

Again to keep the notation simpler, we make an unrealistic assumption about the way interest is paid. We will assume the person does not get paid any interest until the *last day of the year*. What I mean is this: if the person holds B_1 on the first day of the year, the daily return to holding those bonds $\left(B_1 \bullet \dfrac{i}{365} \right)$ is not paid until the last day of the year.

Under this simplifying assumption, the person's total interest income received at the end of the year is:

$$\frac{i}{365}B_1 + \frac{i}{365}B_2 + \frac{i}{365}B_3 + \ldots\ldots + \frac{i}{365}B_{365} = i\left(\frac{B_1 + B_2 + B_3 + \ldots B_{365}}{365} \right) = iB$$

That's going to be handy: total interest income is the interest rate i times *average* bondholding.

Remember that in the Baumol-Tobin model illiquidity takes the form of a cost. On any day that the person sells bonds (or cashes out his overnight loans rather than rolling them over) he must pay a cost F. The person pays the same cost to get his income at the beginning of the year. One can say that F is the cost of engaging in financial transactions on a day.

N is the number of days in the year that the person engages in financial transactions. The total cost of the person's financial transactions over the year is NF.

Getting paid interest is good. Paying the financial transaction cost F is bad. The person wants to maximize his interest income *minus* the total cost of financial transactions NF. I'll call this "profit," denoted Z.

$$Z = iB - NF$$

The person must choose how much bonds *versus* money to hold on each day, taking as given the income Y he receives on the first day of the year and his daily spending $(Y / 365)$. He will make the choice that maximizes Z.

3) Graphs to describe M_d and M, B_d and B.

We will use graphs to describe possible actions the person could take. One graph describes the person's daily bondholding. It has the days of the year (from 1 to 365) on the horizontal axis, and bondholding on that day on the vertical axis. The other graph describes the person's daily moneyholding in the same way.
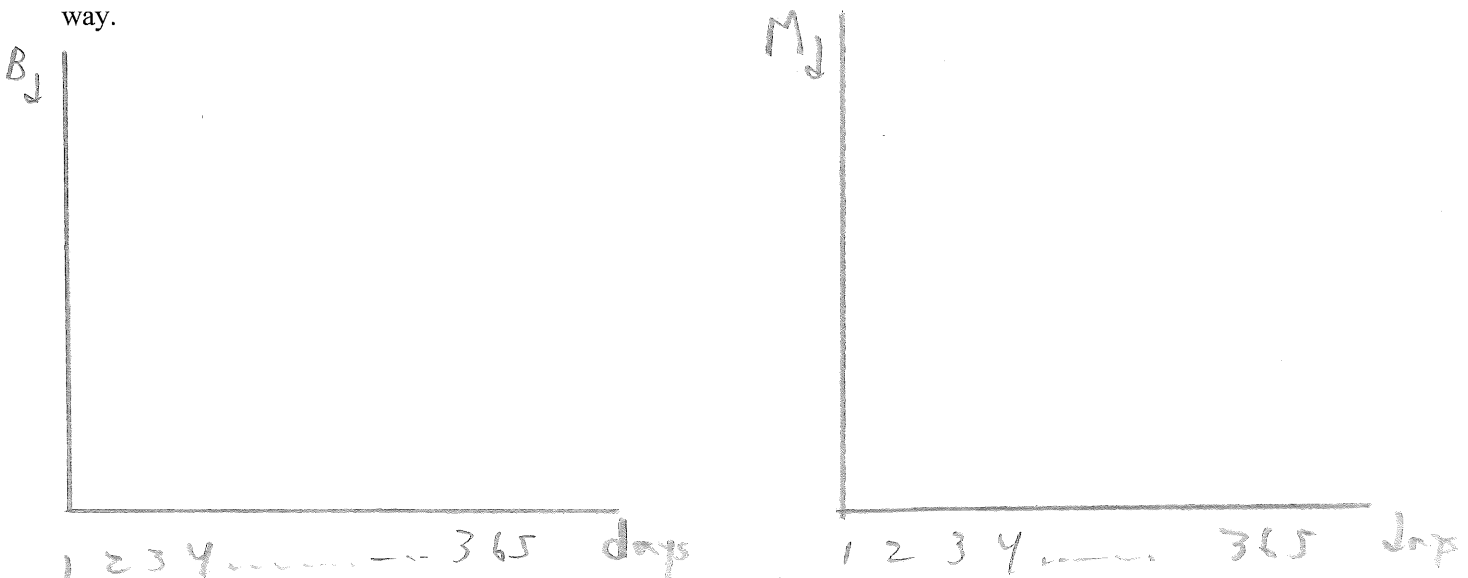
You will need to be able to look at the graphs and figure out the average value of the variable on the vertical axis. Here are some examples to show you how to do it, using M_d as the variable.

Suppose that on the first day of the year M_d is equal to Y. Then it falls steadily until it is zero on the last day of the year. In this case M, the average value of M_d over the year, is equal to $Y/2$.
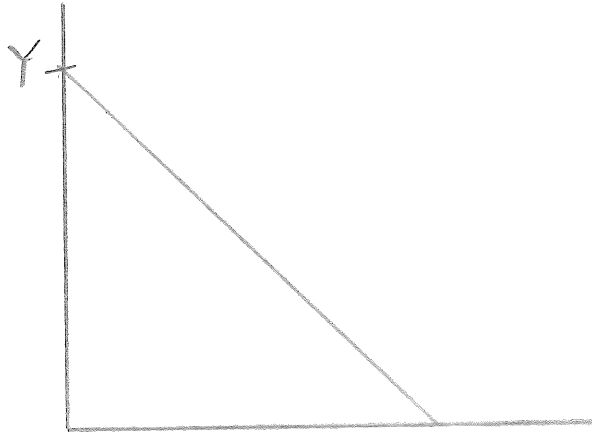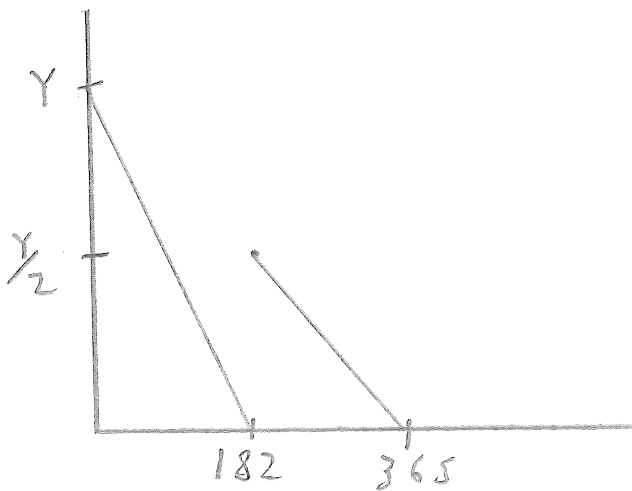
In the case at right $M = Y/4$.
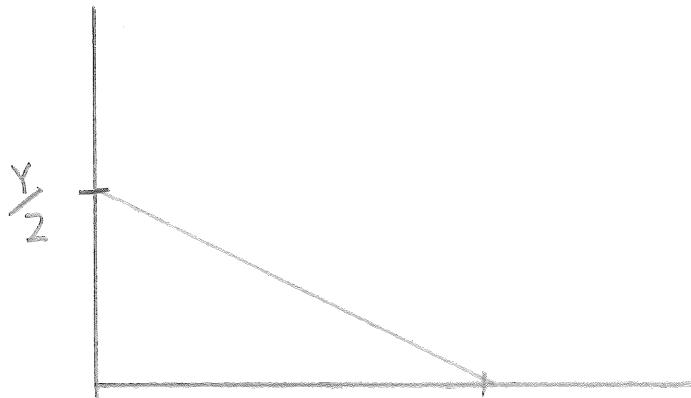
In the case at right $M = \dfrac{3}{8}Y$. This is because for the first half of the year the average value of M_d is $M = Y/2$. For the second half of the year average M_d is $M = Y/4$. For the year as a whole:

$$M = \frac{1}{2}\frac{Y}{2} + \frac{1}{2}\frac{Y}{4} = \frac{2}{8}Y + \frac{1}{8}Y = \frac{3}{8}Y$$

4) Possible actions the person could take

a) Hold just bonds, no money

One thing the person could do is hold no money, just bonds, and sell just enough bonds every day to cover his daily spending. Thus, on the first day he spends his entire income on bonds: $B_1 = Y$. On the second day, he sells enough bonds to pay for his spending, leaving him with $B_2 = Y - \dfrac{Y}{365}$. And so on until the final day of the year, when he sells the last

of his bonds to pay for his last spending of the year. This means he engages in financial transactions on every day: $N = 365$. His daily bondholding and moneyholding are:

B_t

Y

365

M_t

(moneyholding is zero every day)

365
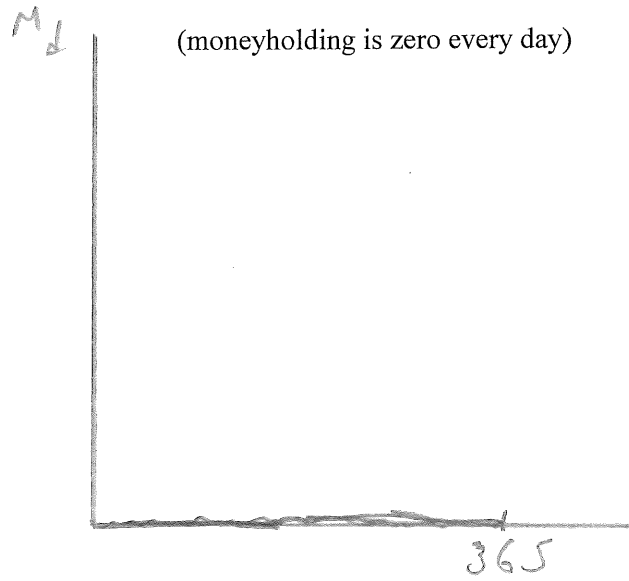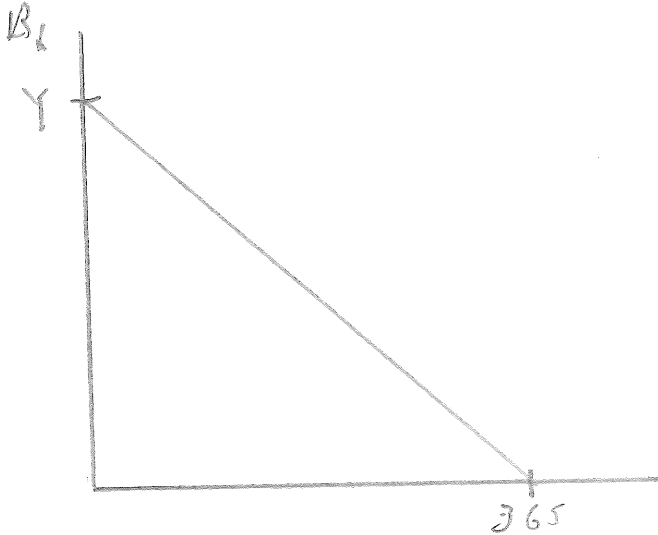
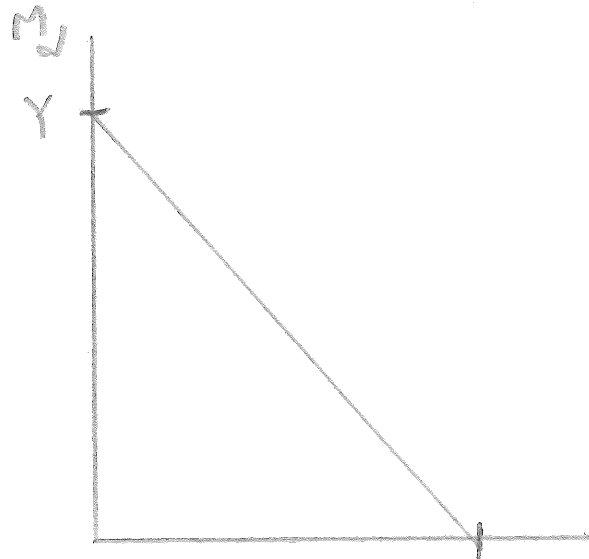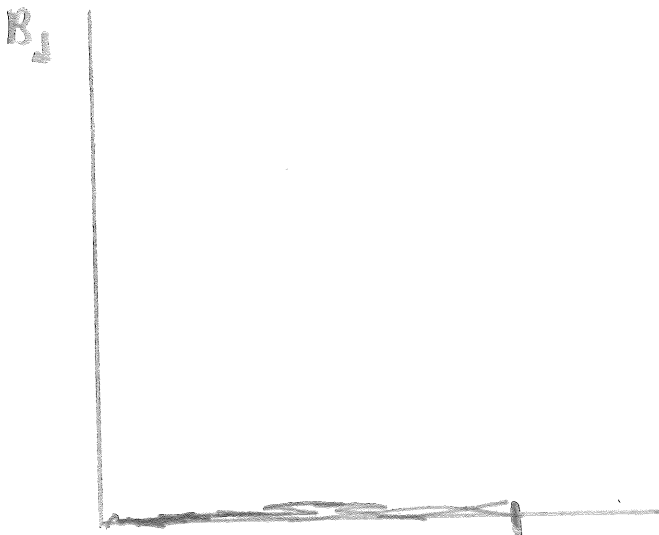His average moneyholding is zero, of course: $M = 0$

What is his average bondholding B? Look at the graph. The average height of that line is $Y/2$.

His interest income is $iB = i\dfrac{Y}{2}$.

His profit is $Z = iB - NF = i\dfrac{Y}{2} - 365F$

b) Hold just money, no bonds

In this case, on the first day of the year the person takes his income in cash, puts the money in a box and takes out just enough money each day to pay for his spending. This means he engages in financial transactions on just one day (when he receives his income): $N = 1$.

B_t

M_t

Y

His average moneyholding is $Y/2$. His average bondholding and interest income are zero, of course.
His profit is zero too: $Z = iB - NF = 0 - 1 \bullet F$

11

c) On the first day, put half his income in bonds, half in money

On the first day of the year the person uses half his income to buy bonds, and puts the other half in money which he puts into a box. Over the first half of the year he spends $(Y/365)$ every day. Halfway through the year he runs out of money. On the day he runs out of money he sells all his bonds. He puts the money he gets into the box. Over the second half of the year he spends $(Y/365)$ every day. At the end of the year he is out of money again. (Then he again receives his annual income Y.) This means he engages in financial transactions on two days: the first day of the year, and the day halfway through the year. $N = 2$.
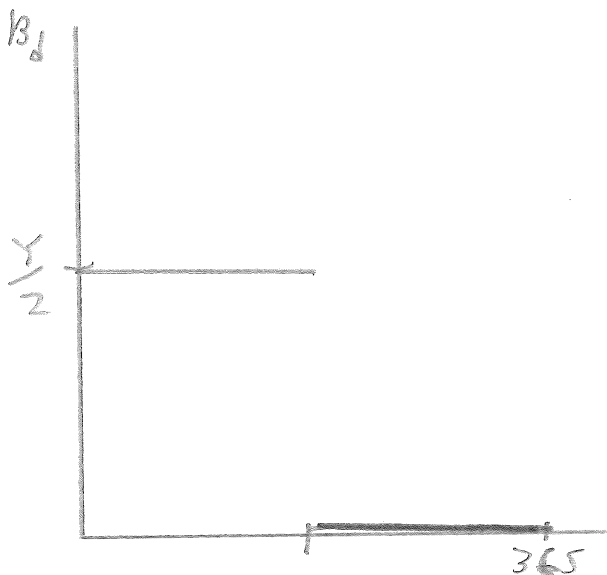
Average bondholding:

$$B = \frac{1}{2}\frac{1}{2}Y + \frac{1}{2} \bullet 0 = \frac{1}{4}Y$$
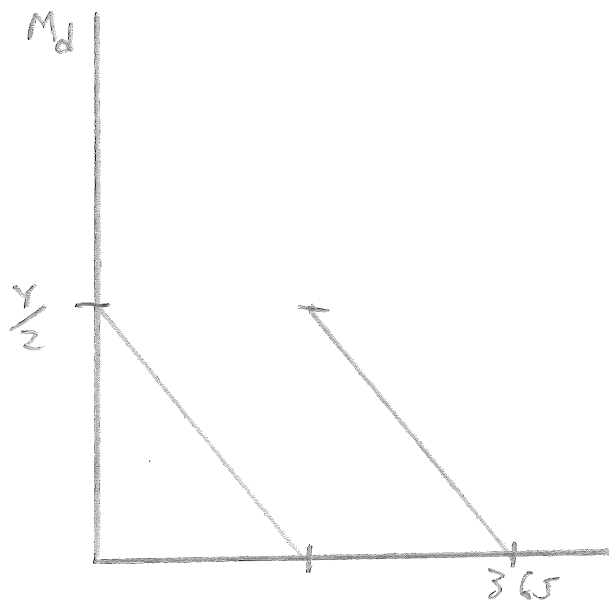
Average moneyholding:

$$M = \frac{1}{2}\frac{1}{2}\frac{Y}{2} + \frac{1}{2}\frac{1}{2}\frac{Y}{2} = \frac{1}{2}\frac{Y}{2} = \frac{1}{4}Y$$

Profit: $Z = iB - FN = i\frac{1}{4}Y - F2$

There are a two things to notice that will be important later on.

First, his average moneyholding is equal to $M = \frac{Y}{2N}$ (because N = 2)
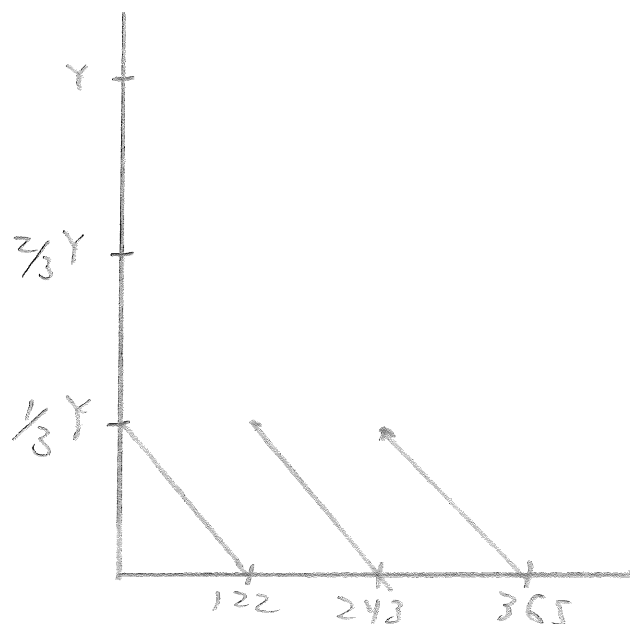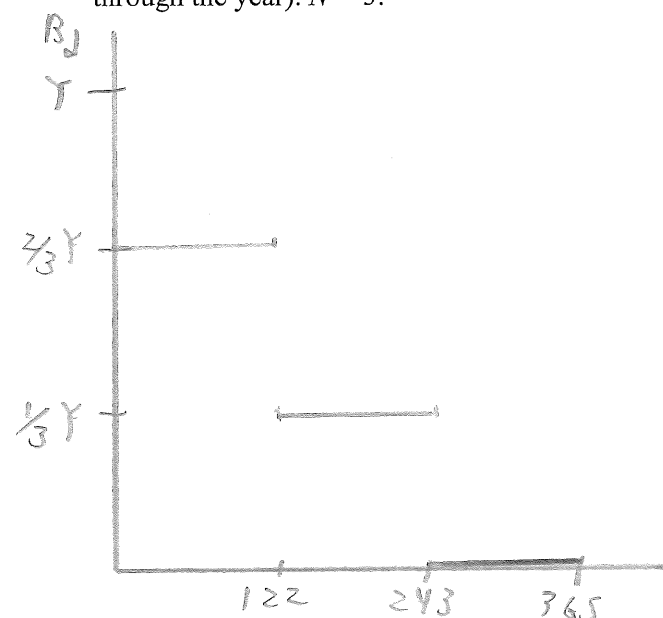
Second, his average bondholding is equal to $\frac{1}{2}Y$ minus his average moneyholding:

$$B = \frac{1}{2}Y - M = \frac{1}{2}Y - \frac{1}{4}Y = \frac{1}{4}Y \qquad \text{so we can describe his profit as:}$$

$$Z = i\left(\frac{1}{2}Y - M\right) - FN = i\frac{1}{2}Y - iM - FN = i\frac{1}{2}Y - i\frac{Y}{2N} - FN$$

d) On the first day, put two-thirds of his income in bonds, one-third in money

In this case, on the first day of the year the person uses two-thirds of his income to buy bonds and puts the remaining third in money, which he puts into a box. Over the first third of the year he spends $(Y/365)$ every day. One-third of the way through the year he runs out of money. The day he runs out of money he sells half of the bonds he bought at the beginning of the year. He puts all the money he gets from selling his bonds into a box. Over the middle third of the year he spends $(Y/365)$ every day. Two-thirds of the way through the year he runs out of money again. He sells the remaining bonds, puyts the money in box and spends it over the rest of the year; it runs out on the last day. This means he engages in financial transactions on three days (the first day of the year, the day one-third through the year and the day two-thirds through the year). $N = 3$.

Average bondholding:

$$B = \frac{1}{3}\frac{2}{3}Y + \frac{1}{3}\frac{1}{3}Y + \frac{1}{3}\cdot 0 = \frac{3}{9}Y = \frac{1}{3}Y$$

Average moneyholding:

$$M = \frac{1}{3}\frac{1}{2}\frac{Y}{3} + \frac{1}{3}\frac{1}{2}\frac{Y}{3} + \frac{1}{3}\frac{1}{2}\frac{Y}{3} = \frac{1}{2}\frac{Y}{3} = \frac{1}{6}Y$$

Profit: $Z = iB - FN = i\frac{1}{3}Y - F3$

Notice that the two things we noticed before are still true here.

His average moneyholding is equal to $M = \dfrac{Y}{2N}$ (because N = 3)

His average bondholding is equal to $\dfrac{1}{2}Y$ minus his average moneyholding:

$$B = \frac{1}{2}Y - M = \frac{1}{2}Y - \frac{1}{6}Y = \frac{3}{6}Y - \frac{1}{6}Y = \frac{2}{6}Y = \frac{1}{3}Y$$

so we can describe his profit as $Z = i\left(\dfrac{1}{2}Y - M\right) - FN = i\dfrac{1}{2}Y - iM - FN = i\dfrac{1}{2}Y - i\dfrac{Y}{2N} - FN$

e) A generalization

We could go on doing this for every possible choice the guy could make. That would be unpleasantly tedious. We can simplify the process some by using the last thing we noticed, that:

$$Z = i\frac{1}{2}Y - i\frac{Y}{2N} - FN$$

This equation holds for all possible choices. Try some more, and you'll see. For example, consider what happens if the guy, on the first day, puts 3/4 of his income in bonds and 1/4 in money.

What we have here is a mathematical function that gives the value of Z. It shows that Z is determined by one variable the guy can control, N, and the three variables the guy takes as given: i, Y and F.

Using this function, we could calculate Z for all the various possible values of N, taking i, Y and F as given. We could see which N maximizes Z. That tells us what the guy will choose to do. It also tells us the guy's average money balance - his money demand - because $M = \dfrac{Y}{2N}$.

We could choose values for i, Y and F and plot out the function on a graph with N on the horizontal axis, Z on the vertical axis. The graph would look like:
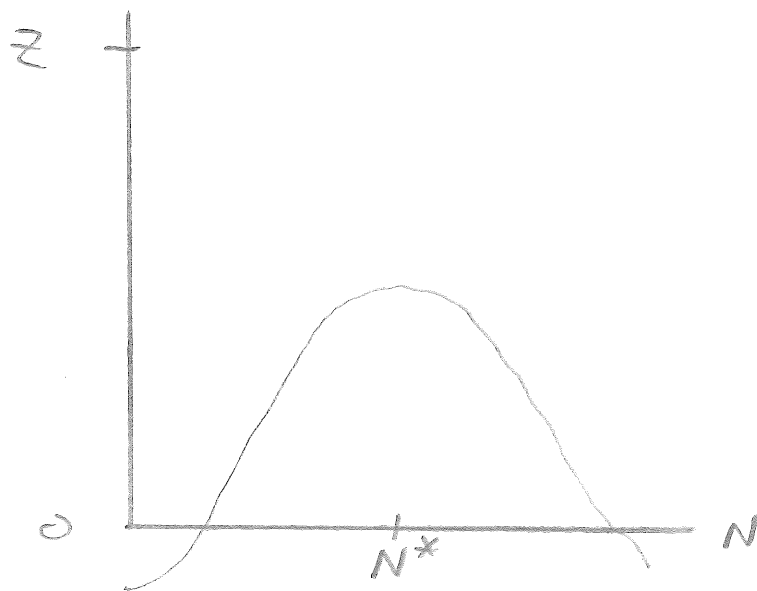
The point on the horizontal axis that I labeled N^* is the optimal value of N. And the person's money demand would be:

$$M^D = \frac{Y}{2N^*}$$

But even this would be tedious. And we would find that N^* depends on the specific values of i, Y and F that we used. We would have to do it over and over again for various combinations of i, Y and F .

Fortunately, there is a shortcut. Calculus!

4) How to find the answer with calculus

a) The method

When you have a mathematical function that gives a relationship between a variable you control (in this case N) and a variable you want to maximize (in this case Z), you can find the optimum value of your "control variable" by taking a derivative and doing some algebra.

Look again at our graph of Z and N. At N^*, the slope of the line is zero. (At the very tip top of a hill, the ground is perfectly level.)

The slope of the line is $\partial Z / \partial N$. Which is to say, when we set $N = N^*$, the value of $\partial Z / \partial N$ must be zero.

So here's the general method. You take your mathematical function. You take the derivative of the function with respect to your control variable - that is you get an equation like:

$$\partial Z / \partial N = \ldots\ldots\ldots$$

Your control variable will be somewhere on the right-hand side of this equation. Using algebra, you see what value of the control variable makes $\partial Z / \partial N$ equal to zero. That's the optimal value.

b) Applying the method here

Step one. Write down the function in a way that will make it easy to take the derivative:

$$Z = i\frac{1}{2}Y \ - i\ \frac{Y}{2N} - FN = i\frac{1}{2}Y - i\frac{Y}{2}\ N^{-1} - FN$$

Step two. Take the derivative:

$$\frac{\partial Z}{\partial N} = -\ i\ \frac{Y}{2}\ (-1)\ N^{-2} - F$$

Step three. Write down an equation that says $\partial Z / \partial N$ equals zero if, and only if, $N = N^*$:

$$0 = -\ i\ \frac{Y}{2}\ (-1)\ N^{*-2} - F$$

Step four. Use algebra to solve the equation for N^*. If you do it right, you'll find:

$$N^* = \sqrt{\frac{iY}{2F}} \qquad \text{There are several other ways to write this, of course, for example } N^* = (iY)^{1/2}(2F)^{-1/2}.$$

Step five. Don't forget step five! Step five is to take the value of N^* that you just derived and plug it into the equation $M = \dfrac{Y}{2N}$. After all, it's M we care about, not N. If you do it right, you'll find:

$$M^D = \frac{Y}{2N^*} = \sqrt{\frac{YF}{2i}}$$

15

The last equation is the money demand function. It gives the person's quantity of money demanded as a function of i, Y and F.

c) Checking the answer

Is the money demand function consistent with

$$(M^D / P) = L(\overset{-}{i}, \overset{+}{Y}) \quad ?$$

Yes. Looking at our answer $M^D = \sqrt{\dfrac{YF}{2i}}$, you can see that $i \uparrow \rightarrow M^D \downarrow$ and $Y \uparrow \rightarrow M^D \uparrow$.

But that's not all. You see that there's an additional term, F, which can affect money demand. The function says $F \uparrow \rightarrow M^D \uparrow$. That is to say, the larger is the cost of engaging in a financial transaction - the less liquid is the interest-paying asset - the more money the person wants to hold.

5) Another way to set up the problem: minimizing cost

We said that the person is choosing N to maximize a function:

$$Z = i\frac{1}{2}Y - i\frac{Y}{2}N^{-1} - FN$$

This function has three parts.

Part one is $i\dfrac{1}{2}Y$. This is the interest income the person would receive if he held no money at all. Notice that this part of the function is not affected by N.

Part two is $-i\dfrac{Y}{2}N^{-1}$. This is the opportunity cost (lost interest) of the average moneyholding. It is affected by N.

Part three is $-FN$. This is the cost of the person's financial transactions. It is also affected by N.

Let's add the two costs together to get a "total cost" function $TC = i\dfrac{Y}{2}N^{-1} + FN$.

Then we can say $Z = i\dfrac{1}{2}Y - TC$

Look! The value of N that *maximizes* Z must be the value that *minimizes* TC, because $i\dfrac{1}{2}Y$ is not affected by N.

To find the value of N that minimizes TC we could plot out the total cost function on a graph that has N on the horizontal axis, TC on the vertical axis. The graph looks like:

Cost

$$TC = i\frac{Y}{2N} + FN$$

FN

$$i\frac{Y}{2N}$$

N
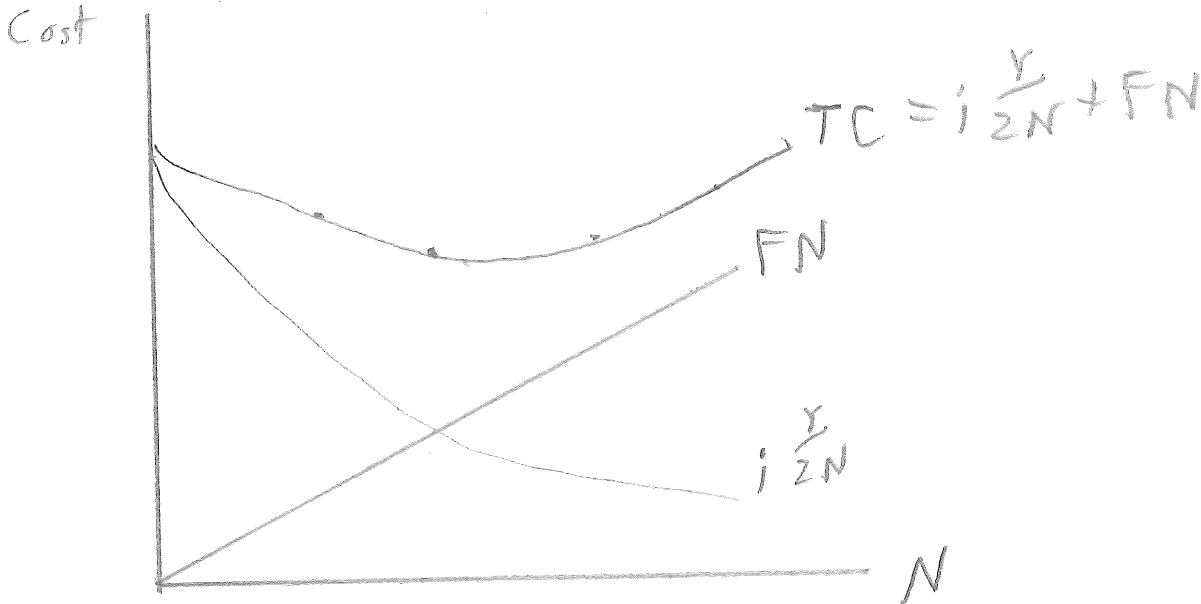
Notice that I have put three lines on the graph. One is the opportunity cost of moneyholding. One is the cost of financial transactions. Taking the vertical sum of these two lines gives the TC line. N^* is the value of N at the bottom of the TC line.

We can use our calculus trick here too, because the slope of the TC line is zero at the minimum. (At the very bottom of a valley, just like at the top of a hill, the ground is perfectly flat.)

Step one. Write down the TC function in a way that will make it easy to take the derivative:

$$TC = i\,\frac{Y}{2N} + FN = \; i\frac{Y}{2}\,N^{-1} + FN$$

Step two. Take the derivative:

$$\frac{\partial TC}{\partial N} = \; i\frac{Y}{2}\,(-1)\,N^{-2} + F$$

Step three. Write down an equation that says $\partial TC / \partial N$ equals zero if, and only if, $N = N^*$:

$$0 = \; i\frac{Y}{2}\,(-1)\,N^{*-2} + F$$

Step four. Use algebra to solve the equation for N^*. If you do it right, you'll get the same answer we got before:

$$N^* = \sqrt{\frac{iY}{2F}}$$

And step five (don't forget it!). Plug $N*$ into the equation $M = \dfrac{Y}{2N}$:

$$M^D = \sqrt{\frac{YF}{2i}}$$

Same answer as before, of course.

This second way to find the money demand function, by minimizing TC rather than by maximizing Z, is the way that *most* people describe the Baumol-Tobin model. It is a little easier than maximizing Z because you don't have to keep track of $i\dfrac{1}{2}Y$.

I suggest that you, too, use the minimizing-TC method. But don't forget that it is equivalent to maximizing "profit" Z.

H) How is our economy different from the one we have examined in this section?

In the introduction to this section, I said that the economy described by the liquidity-preference money demand graph from Econ 362, and the specific Baumol-Tobin version of the model, is different from our economy. How is it different?

People in our economy usually do not pay for things with money. Sometimes people pay for things with cash, sure. But usually people pay for things with debit cards, checks, or credit cards. (Young people also use Venmo.) All of those forms of payment involve *banks*.

Also, some of our money, the ledger money held by banks at the central bank, pays interest.

You will learn about banks in the next section.

©**Christopher Hanes**   **Economics 450   Class notes**

IX) Financial intermediation
    A) Intro
    B) How financial intermediaries make profit
        1) In general
        2) Three strategies
            a) Profiting from default risk premiums
            b) Profiting from term premiums
            c) Profiting from liquidity premiums
        3) Combination strategies
    C) Profit *versus* net worth of a financial intermediary
    D) Banks
        1) Introduction
        2) Demand deposits
        3) Time deposits
        4) Checkable deposits
            a) Definition
            b) What happens when you make payments with a checkable deposit
            c) Reserves
            d) Interest rates on checkable deposits can be very low
    E) How things can go wrong
        1) Introduction
        2) The risk of profiting from default-risk premiums: lots of defaults at once
        3) The risk of profiting from term premiums: interest-rate risk
        4) The risk of profiting from liquidity premiums: fire sales
        5) Runs and multiple equilibria with self-fulfilling expectations
            a) Introduction
            b) First come, first served
            c) Simple example
            d) Conclusion
    F) Capital
        1) Introduction
        2) Starting a new financial intermediary
        3) Continued operation of the FI
            a) In good years: add more capital?
            b) In bad years: recapitalize?
    G) Regulation of financial intermediaries
        1) Introduction
        2) How FIs would operate in the absence of regulation
        3) Possible negative externalities of a FI's bankruptcy
            a) Contagion
                i) Introduction
                ii) Networks of lending among FIs
                iii) Sales of illiquid assets
                iv) Simple example
            b) Macroeconomic effects
                i) Introduction
                ii) Reduction in supply of loans to businesses and households: "credit crunch"
                iii) Financial crises that affect banks: no way to make payments
        4) Regulations applied to FIs
            a) Introduction
            b) Capital requirements

> c) Liquidity requirements
> d) Stress tests
> e) Deposit insurance
> f) Special bankruptcy procedures for FIs

IX) Financial intermediation

A) Intro

Businesses borrow money for many reasons (e.g. to purchase new buildings and equipment, to buy inventories of goods to sell, to buy other businesses). In this section I discuss businesses that borrow money *in order to lend it on to other entities* - to households, or to governments, or to other businesses. A business that borrows in order to lend to others is called a "financial intermediary" which I will abbreviate as FI. What a FI does is called "financial intermediation." A FI lends by buying bonds, making loans, or both. Examples of FIs include banks, investment banks, hedge funds and many types of specialist lending companies such as mortgage companies.

The bonds that a FI buys and the loans it makes are assets on its balance sheet. I say that a FI "acquires assets" by buying bonds or making loans. At the same time that the FI is making loans and buying bonds it is *taking* loans and/or *issuing* bonds. The loans and bonds through which the FI borrows are liabilities on the FI's balance sheet.

The business of financial intermediation is inherently risky. A financial intermediary cannot earn profit without taking on risk. Most businesses have an element of risk, of course. But financial intermediation is somewhat special in that the failure of a firm engaging in financial intermediation can harm a great many other firms and people. Thus, financial intermediaries are among the most heavily regulated types of businesses.

B) How financial intermediaries make profit

1) In general

A FI simultaneously borrows and lends. It makes a profit if the returns on its lending are higher than the interest rates it pays on its borrowing. Thus a FI looks for and exploits spreads in interest rates or rates of return. As you have learned, such spreads are due to default-risk premiums, liquidity premiums and term premiums.

2) Three strategies to make profit

a) Profiting from default risk premiums

A FI can profit by borrowing money to acquire default-risky bonds and loans, if people think that the FI *itself* is unlikely to default on the loans it takes and/or the bonds it issues. Thus returns on the FI's assets are boosted by default-risk premiums, while interest rates on the FI's borrowing are lower because they contain smaller default-risk premiums.

How can there be low risk of default on lending *to* the FI if there is high risk of default on the bonds and loans the FI holds? This can be true if the FI acquires a *lot* of default-risky assets and it is unlikely that a large fraction of them will default *at once*. A high probability of default on an individual asset is not a problem as long as the particular events that could cause default on one asset are uncorrelated with the events that cause defaults on the others.

b) Profiting from term premiums

A FI can profit by borrowing overnight or issuing short-term bonds, and using the proceeds to buy long-term bonds. Because of term premiums, yields on long-term bonds are usually higher than yields on short-term bonds.

c) Profiting from liquidity premiums

A FI can profit by borrowing overnight or issuing liquid bonds and using the proceeds to buy illiquid bonds and/or make loans. The returns on the FI's assets will be high due to liquidity premiums. The yields on the bonds *issued* by the FI will be low as long as the bonds issued by the FI are liquid. The bonds issued by an FI will be liquid if information about the risk of default by the FI itself is sufficiently symmetric.

To make loans, a FI can use the techniques described in the previous section (lend only to borrowers with enough net worth; collateral; screening; monitoring and restrictive convenants).

3) Combination strategies

For simplicity I have described the three strategies as if they were distinct but most things FIs do are combinations of the strategies. For example, many FIs borrow overnight and use the proceeds to buy long-term illiquid bonds issued by companies and make long-term loans to companies. The interest rates on the FI's overnight borrowing can be low because an overnight loan is a perfectly liquid asset to the lender: to cash out, the lender just refrains from rolling over the loan. Meanwhile returns to the long-term corporate loans and bonds held by the FI are high because of term premiums, liquidity premiums and default-risk premiums, all at once.

C) Profit *versus* net worth of a financial intermediary

The annual profit that a FI makes in a year is the earnings on its *assets* minus the interest it has to pay on its borrowing. Earnings on its assets is:
> interest received on the bonds and loans held by the FI
> + gains from buying assets and selling them before maturity at a higher price
> - losses from buying assets and selling before maturity at a lower price

The owners of a FI care about its annual profit. But there is another number that matters too. That is the financial intermediary's *net worth*. I talked about net worth earlier, in section VII). An entity's net worth is the current market value of its assets *minus* the current market value of its liabilities. An entity is *"solvent"* if its net worth is positive. It is *"insolvent"* if its net worth is negative.

For a financial intermediary, net worth is the current market value of the bonds, loans and any other assets that it holds *minus* the current market value of its borrowings. A FI's net worth is *not the same* as its annual profit. As an example, consider a FI that borrows $1 million overnight (rolling over the borrowing, day after day) and uses the borrowed funds to buy long-term coupon bonds. Its annual profit is:
> coupon payments on the long-term bonds in that year
> + any increase over the year in the market price of those long term bonds
> - any decrease over the year in the market price of those bonds
> *minus* the total interest paid over the year on the $1 million of rolled-over borrowing.

This FI's net worth, on the other hand, would be calculated as of a particular day in the year. Its net worth on that day would be the market value of the bonds on that day *minus* $1 million (the amount of its borrowing on that day).

It is possible for a FI to be insolvent even if its current profit is positive. Go back to my example. Suppose long-term bond yields increase so much that the market value of the FI's bonds falls below $1 million. Then the FI is insolvent. But the increase in yields will not change the coupon payments on the bonds. They may still be enough to pay the interest on the FI's overnight borrowing.

Generally, bad things can happen if a FI becomes insolvent. This is true even if the FI's current earnings on its assets are enough to pay the current interest on its borrowing. I will return to this point below.

D) Banks

1) Introduction

A bank is a particular type of financial intermediary. A typical bank engages in a variety of FI strategies. It borrows partly by taking short-term loans, including overnight loans, and partly by issuing bonds at both long and short maturities. On the asset side, a bank buys bonds and makes loans of all maturities. Because it does a lot of lending, it hires a lot of people to screen and monitor borrowers. But these things are not the distiguishing features of banks.

The distinguishing feature of a bank is an additional way that it borrows, in addition to issuing bonds and taking loans. A bank also takes *deposits*. In fact, taking deposits is usually a bank's *main* way of borrowing. There are various types of deposits.

2) Demand deposits

To borrow by taking demand deposits, a bank makes this offer to the public: "give me some money, and I'll give it back to you whenever you ask for it, with interest compounded daily at an annual rate of x percent." A person who lends money to the bank on these terms is a *depositor*. The money he lends to the bank on these terms is called a *demand deposit*, because the bank has promised to give him his money back on demand.

In effect, a demand deposit is an overnight loan to the bank. But it has special characteristics. It "rolls over" automatically until the lender withdraws the deposit. While the interest rate on an ordinary overnight loan is negotiated daily between borrower and lender, the interest rate on a demand deposit is set by the bank and changed somewhat infrequently.

3) Time deposits

Some special types of deposits, called "time deposits," have somewhat different rules. To withdraw a time deposit the depositor must give the bank a certain amount of advance notice (like a week or two), or must wait until after a certain pre-specified date. Most deposits are demand deposits, not time deposits. When I say "deposit" I usually mean "demand deposit."

4) Checkable deposits

a) Definition

A checkable deposit is a demand deposit with an extra feature: the depositor can order the bank to pay a specified amount of the money the bank owes the depositor to a third party. Originally such payment orders were given in the form of notes written on paper. These notes are called "checks," hence "checkable deposits." Nowadays there are other ways for a depositor to communicate payment orders to a bank, such as debit card networks.

An ordinary savings account is a demand deposit, but not a checkable deposit. A checking account *is* a checkable deposit, but it is not the only kind of checkable deposit. Some other types of deposits (e.g. "money market mutual funds") are not called checking accounts but they are checkable deposits on our definition, because you can use them to make payments.

b) What happens when you make payments with a checkable deposit

When you write a check, the payee (the person to whom you made out the check) can, if she wants, take the check to your bank and ask for the payment in cash.

Most people don't do this, of course. Most people have bank accounts. If you make out a check to someone with a bank account, that person will probably take the check to her own bank and "deposit" it into her own account. Most types of payment order other than checks, such as those made through debit card networks, *must* be received

through the payee's bank - the payee does not have the option to ask for money directly. When a payee deposits a check into her bank account she is asking her bank to:

- collect the money *for her* from the payer's bank and then
- add that amount to her checkable deposit.

If the payer and payee have accounts at the same bank, the payer's bank doesn't actually hand over money to anyone. The bank just debits the amount of the check from the payer's account and credits the same amount to the payee's account.

But most of the time the payee's account is at a different bank. In that case, the payer's bank *does* have to hand over money, to the payee's bank.

c) Reserves

Every day a bank has to pay out lots of money to other banks to cover its depositors' payment orders, and to payees who ask for direct payment in cash. A bank also has to pay money to depositors who ask for some or all of their deposits back, that is depositors who withdraw cash from their accounts. At the same time, a bank is taking in money from other banks as a result of the payment orders received by its deposits, and new deposits of money.

On some days the inflow of money is greater than the outflow. On some days the inflow and outflow balance. But on many days the outflow is bigger than the inflow, so the bank has to pay out money on net.

How does a bank make sure it can always pay out money as needed? Mostly, by making sure that a large enough portion of its assets are the kind of assets that can be "liquidated" - turned into cash - within a day. This means Treasury bonds and overnight loans - I mean overnight loans the bank makes, not the overnight loans the bank takes (which are a liability). Overnight loans are perfectly liquid: a lender liquidates an overnight loan simply by not renewing it.

But banks also hold some money, plain old money as defined in section VIII. When a bank holds money, just money, the money it holds is called "reserves." In a country without a central bank, such as the United States prior to 1914 (when the Federal Reserve system was founded), a bank holds money in the form of coins and paper money in its vaults. In a country with a central bank, such as the United States today, most of the money held by banks is in the form of ledger money at the central bank, that is numbers in the central bank's account books. This type of central bank ledger money is called "reserve accounts." I'll get back to this in a later section of the course.

d) Interest rates on checkable deposits can be very low

Making payments through checkable deposits is convenient for depositors. It is a service valued by depositors. Thus, banks can get away with paying very low interest rates on checkable deposits - sometimes no interest at all. From the bank's point of view, the fact that it can get away with paying very low interest rates on checkable deposits is a good thing, but this is partly counterbalanced by costs to a bank of clearing checks, operating debit card networks and so on.

E) How things can go wrong

1) Introduction

Recall that a financial intermediary makes a profit only if the returns on its lending are higher than the interest it pays on its borrowing; thus, it must acquire assets whose returns are boosted by premiums due to default risk, interest-rate risk (term premiums) or illiquidity (liquidity premiums). This makes the business of financial intermediation inherently risky.

2) The risk of profiting from default-risk premiums: lots of defaults at once

I said that a FI can profit from acquiring default-risky bonds and loans as long as there is little risk that an extraordinarily large fraction of the FI's assets will go into default *at once*. The danger is that a lot of the FI's assets *do* go into default at once. This can happen if there is some extraordinary, bad event that affects many of the FI's borrowers at the same time so they all default at once. For an example, consider a bank that lends mainly to local businesses and households in a county. Suppose a hurricane destroys businesses in the county, so that those businesses can't repay their business loans, and people who used to work for the businesses can't repay their mortgage and car loans. They default *en masse*.

3) The risk of profiting from term premiums: interest-rate (duration) risk

Suppose that a FI follows a "term premium" strategy: it borrows short-term in order to borrow default-free, long-term Treasury bonds, paying the interest on its borrowing with the coupon payments on the bonds. Treasury bonds are default-free. So what's dangerous about this? The danger is interest-rate risk. A big increase in bond yields might reduce the bonds' market values enough to make the FI insolvent.

Now, as noted above, the FI may still be able to pay the interest on its short-term borrowing and earn an annual profit, because the drop in the bonds' market value does not affect their coupon payments. But the FI's insolvency may be a problem nonetheless. Suppose that there is a big increase in bond yields *and*, at the same time, a lot of the short-term lenders to the FI suddenly withdraw their loans - ask for their money back. Then the FI will have to try to pay off the loans by selling its long-term bonds. But because the current market price of the bonds is low, selling the bonds might not generate enough funds to pay off all the withdrawn loans. If this is the case, the FI will have to declare it cannot pay off its debts - it is bankrupt.

When lenders to a FI withdraw their loans *en masse*, that is called a "run" on the FI. In the case of a FI that has been taking ordinary short-term loans, it means lenders stop rolling over their loans. If the FI has been selling short-term bonds, it means that people are unwilling to buy the FI's next issue of short-term bonds. In the case of a bank, a run occurs when many depositors withdraw their deposits all at once. Before ATMs, to withdraw a deposit from a bank you had to go to a branch of the bank and get your money from a teller. Before cars, you had to get to the bank branch at least partly on foot. Hence the term "run": if you were in a hurry to get your money, you would *run* to the bank.

Note that the problem is the *combination* of a run and an increase in bond yields. In the absence of the run, the FI would have been insolvent but it would still have been able to pay interest on the loans to the FI and keep operating.

4) The risk of profiting from liquidity premiums: fire sales

Now consider an FI that follows the strategy of buying illiquid assets to profit from liquidity premiums. Perhaps it plans to hold the assets until they pay off (until bonds pay off their IOUs or borrowers pay off their loans). Or perhaps to plans to sell the assets before they mature, gradually, with enough time to find knowledgable buyers willing to pay relatively high prices for the illiquid assets.

The danger is that for some reason there will be a run on the FI. To pay off the withdrawn loans, the FI will have to sell off its illiquid assets quickly. That means fire sales: low prices. With the assets valued at these low fire-sale prices, the FI may be insolvent. If so, selling the FI's assets will not generate enough funds to pay off all the withdrawn loans. The FI will have to declare it cannot pay off its debts - it is bankrupt.

An event like this is called a "liquidity crisis." It may sound like the interest-rate risk problem I described above, but it is different. In the interest-rate risk case, the FI is insolvent whether or not there is a run. In a liquidity crisis the FI is insolvent only because there was a run on the FI, forcing a fire sale. If there had been no run, the FI would not be insolvent. The tricky thing is that illiquid assets do not have a "market price," in the same way that Treasury

bonds do. The value of an illiquid asset is one number if it can be sold to a knowledgable buyer, a lower number if it must be sold in a fire sale.

5) Runs and multiple equilibria with self-fullfilling expectations

a) Introduction

Now let's turn to the question: why might a run occur? The answer is that, under an additional condition, a run can occur easily, for very little reason - perhaps for no reason at all. If lenders think that lenders will keep lending to the FI, then lenders *will* keep lending to the FI, with good results. If lenders think that lenders will stop lending to the FI, then lenders *will* stop lending to the FI, with bad results. Economists call a situation like this "multiple equilibria" with "self-fulfilling expectations."

b) First come, first served

What is the "additional condition"? It is called "first come, first served." In this context, "first come, first served" means that the short-term lenders to the FI who ask for their money back *first* are most likely to be repaid, while lenders to the FI who are slower to ask for their money back are likely to get less, maybe nothing. Consequently a lender to the FI who is deciding whether to keep lending to the FI must guess that *other* lenders to the FI will do that day and try to do the same thing. If she thinks that other lenders will keep lending to the FI, it is safe to keep lending to the FI herself. But if she thinks other lenders are about to withdraw their loans to the FI, she should withdraw her own loan to the FI as fast as possible. Thus, the cause of a mass withdrawal of short-term lending to an FI may be a belief on the part of lenders that *other* lenders are about to withdraw.

c) Simple example

Here is an example. It is a "model" of a run with multiple equilibria. To keep things simple, we will assume there are just two lenders to the FI, and we will consider just one day.

Suppose there is an FI that would be in trouble if a run occurs, either because it is holding illiquid assets, or because it is holding liquid long-term assets whose market value has been depressed by an increase in yields. The short-term loans to the FI have been *overnight* loans. The two lenders to the FI are called Bob and Jane. Each has been lending $X to the FI. The FI has promised to pay an overnight interest rate denoted *i*. On the morning of the day in question, Bob must decide whether to withdraw his loan or roll it over for another day. So must Jane. We assume that Bob and Jane want to maximize the expected values of their investments.

Four things can happen:
1. Bob and Jane both roll over their loans.
2. Bob withdraws, Jane rolls over.
3. Jane withdraws, Bob rolls over.
4. Bob and Jane both withdraw.

What happens in each case?

1. Bob and Jane both roll over their loans.
If Bob and Jane both roll over their loans to the FI, then the FI can pay both of them the interest it promised. Each of them gets $(1+i)X$.

2. Bob withdraws, Jane rolls over   or   3. Jane withdraws, Bob rolls over.
If either withdraws, the FI must sell its assets suddenly. At the low prices that the assets will fetch (because they are illiquid assets in a fire sale, or because their market value has been depressed by an increase in yields), there will not be enough money to pay off both Bob and Jane. In fact, the sale will generate only $X - enough to pay off just

one of the lenders. Thus, the one who withdraws gets paid. The one who did not withdraw is left holding the bag. This is "first come, first served."

4. Bob and Jane both withdraw.
Again the FI must sell its assets and gets only enough money to pay out $X. Who gets paid, Bob or Jane? Assume that they are equally fast in running to ask for their money back, so that they are equally likely to be the first one who asks. The probability that Jane is first is 1/2. The probability that Bob is first is 1/2. The expected value of the money one of them receives is $\frac{1}{2}X + \frac{1}{2}0 = \frac{1}{2}X$

We can describe the situation with a **graphic** called a two-by-two matrix. It represents the two decisions each "player" can take, and the four possible things that can happen with four boxes. Each box represents a possible outcome. Inside each box is written the amount of money each "player" would receive in that outcome. You use the graphic to figure out what will happen when the two players, Bob and Janes, "play the game." What the graphic shows is that outcomes 1. and 4. are both possible, with self-fulling expectations. In class, I will show you how to use the graphic. Before class, make sure that you take a good look at it. Here it is:

d) Conclusion

What is the lesson of this model? It is that the common FI strategy of borrowing short-term to acquire illiquid assets is fragile. An event that creates fear that an FI may be insolvent, or fear that *other* people will fear that the FI is insolvent, can cause a liquidity crisis. (Read that last sentence again and make sure you understand it.)

For the same reason, a FI that is holding liquid long-term assets whose market value has been depressed by an increase in yields is also in a fragile situation, even if current earnings on its assets are enough to pay the interest on its short-term borrowing.

F) Capital

1) Introduction

I have said that all three strategies of financial intermediation involve a risk that the FI will be unable to pay off its borrowings. That is bad not only for the owners of the FI but also bad for the lenders to the FI - they don't get their money back! So why would anyone lend to a FI in the first place?

In fact, no one *will* lend to an FI unless the FI has "capital." "Capital" is a thing that decreases the probability that a FI will be unable to pay back the money that it borrows. The more capital that a FI holds, the less likely that it will go bankrupt. A FI cannot borrow at all unless it has enough capital.

So what is capital? From one point of view, capital is just the net worth of the FI. It is the value of the FI's assets (the bonds it has purchased, the loans it has made) minus its liabilities (the loans it has taken, the bonds it has issued). From another point of view, capital is money that the owners of a FI have put into the business. It is money that they have given to the FI so that the FI can acquire more assets. Without capital, the value of the FI's assets would be equal to the amount of its borrowing, at least at first. With capital, the value of the FI's assets is greater than its borrowing.

It might sound like capital is a loan to the FI from the owners, but it is not. An FI's liabilities are *just* its borrowing *not* including its capital. The owners do not earn interest on their contribution of capital. They do not receive, in return for contributing capital, promises of future IOU's as with a bond. The owners do receive the profits of the FI. But they receive the FI's profits anyway. Contributing capital to the FI does not increase the owners' profits in any direct way. (It may increase the owners' profits indirectly as I will explain below.) But an FI without capital cannot borrow.

The easiest way to understand the role of capital in an FI is to imagine starting a new FI. So we'll begin there.

2) Starting a new financial intermediary

Suppose you want to go into the business of financial intermediation. You plan to sell some bonds or take some loans, paying relatively low interest rates on your borrowing. You will take the money you borrowed and buy bonds and/or make loans following one or all of the three strategies described above.

From your point of view, this seems like a no-lose proposition. Let's say you borrow $1 million one way or the other and acquire $1 million of long-term bonds, illiquid bonds and long-term loans. If everything goes well, you get to keep the difference between the returns on your assets and the cost of your borrowing. If things go badly, the FI becomes insolvent, goes bankrupt, and you get nothing. But you lose nothing. Thus, there is some probability you will get nothing, but a positive probability you will get something. The expected value of the business is positive.

If you try to start a FI this way, however, it won't work. No one will lend to you. No one will buy the bonds you issue.

Why? Partly just because of the math of the situation. If your FI's liabilities, your borrowing, is equal to $1 million, and your FI is holding just $1 million in assets - the bonds you bought with the $1 you borrowed - then a very small drop in the value of the assets makes the FI insolvent. If the FI becomes insolvent, it may default on at least some of its borrowing

There is also the reason I described in the section in section VII), "Loans," under "General problem: borrowers have incentive to do high-risk, high-return things...Lend only to borrowers with something to lose." Everyone knows that you have nothing to lose should the FI become insolvent. Everyone knows this gives you incentive to do high-risk, high-return things. You might acquire assets that will pay high returns if everything goes well, but also

have relatively high default risk, interest-rate risk or illiquidity. That will raise the probability that your FI will become insolvent and default on at least part of its borrowing.

So how can you get your FI business started? You must take some large amount of your own money and put it into the FI. That money is capital. With that extra money the FI acquires more in assets, while the FI's liabilities are still $1 million. Now people may be willing to lend to you. First, just because of the math. Since the value of the FI's assets is greater than its liabilities right from the start the FI is no longer on the edge of insolvency. The value of its assets can fall some, and the FI will still be solvent up to a point. Suppose that assets values fall down to the point where assets are just equal to the FI's borrowing. Then you, the owner, have lost your capital. But the FI is still solvent and can still pay off its borrowing.

That brings up the second reason people may be willing to lend to you. They know you have some skin in the game, something to lose if the FI becomes bankrupt. That will discourage you from adopting high-risk, high-return strategies.

Notice that this won't work if the $500,000 is a *loan* from you, or if you buy $500,000 worth of bonds newly issued by the FI. That will allow the FI to increase its assets by $500,000. But the FI's liabilities also increase by $500,000. It is a wash in terms of capital. It won't reduce the probability that the FI becomes insolvent. You have to *give* the money to the FI.

Suppose you succeed in getting the FI started this way. What happens in subsequent years?

3) Continued operation of the FI

a) In good years: add more capital?

In good years there will be no waves of defaults or need for fire sales or big declines in prices of the long-term bonds held by the FI. There will be profit from the difference between the annual return on the assets and the cost of the FI's borrowing. The profit belongs to the FI's owner. That's you. What will you do with the profit?

One thing you can do is pay it all out to yourself as dividends.

Another thing you can do is pay yourself just some (or even none) of the profit, and use the money you don't pay yourself to buy more assets for the FI. When a business does not pay out all of its profits to its owners, the profit kept in the business is called "retained earnings." Retained earnings add to a FI's capital.

Is it a good idea to add to capital by retaining earnings? It depends. There is a possible benefit; there is also a cost.

The cost is that adding to capital tends to reduce the rate of return you receive on your investment in the FI. Your investment is the capital you have contributed. The rate of return on the FI's capital is the FI's profit divided by the amount of capital. For any given cost of borrowing by the FI, adding more capital lowers this ratio. To see this, let i^B denote the average rate (interest rate or yield) on the FI's borrowing. i^A denotes the average rate of earning on the FI's assets. B is the amount of borrowing by the FI, in any form, and C is the capital you put in. Then the annual profit earning by the FI is $i^A(C+B)-i^B B$.

The annual return on the capital is the ratio of annual profit to capital, which is:

$$\frac{i^A(C+B)-i^B B}{C} = i^A + (i^A - i^B)\frac{B}{C}$$

See that an increase in capital C tends to reduce the annual return on the capital you put in.

What is the possible benefit? Increasing capital might make people willing to lend to you at lower rates of interest. In the equation above, that means a decrease in i^A, which tends to *increase* the annual return on capital. Why might an increase in capital decrease i^A? Recall that an increase in capital reduces the probability that the FI will become insolvent on the future. So people might be willing to lend to you at a lower interest rate.

So you have a tradeoff to consider. Presumably you will choose the amount of capital that maximizes the return on your investment of capital.

b) In bad years: recapitalize?

In bad years there are waves of defaults or need for fire sales or declines in prices of the long-term bonds held by the FI. The value of the FI's assets falls, so there is a decrease in its net worth - in its capital. The FI can become insolvent this way, and be closed down by regulators or go bankrupt. But even if the FI does not become insolvent, the decrease in capital can be a problem. It raises the probability that the FI will become insolvent in the future. That might make people unwilling to lend to the FI. Even if people are still willing to lend to the FI, they might demand higher rates of interest in their loans to the FI or higher yields on bonds issued by the FI.

What can you do to make people more willing to lend to the FI? You need to top up the FI's capital. That is called "recapitalizing" an FI. If people are still willing to lend to you, you may be able to recapitalize slowly, over time, by retaining future earnings. But you may need to recapitalize quickly. You can do that by putting in more of your own money.

What if you don't have enough of your own money to bring capital back to where it needs to be to make people willing to lend to you? In that case you will need to get capital from other people. But how can you persuade them to contribute capital to your FI? Remember, you cannot offer to pay interest or future IOU's in return for capital. That would be a loan. It would increase liabilities as well as assets, making it a wash in terms of capital.

What you must do is make them part owners, entitled to a share of future profits (if there are any). That will, of course, reduce *your* share of the profits. Ouch.

G) Regulation of financial intermediaries

1) Introduction

At the beginning of this section I said that financial intermediaries are subject to relatively heavy government regulation. Most of the regulations are intended to reduce the probability that a FI will go bankrupt, by forcing a FI to operate in less risky ways. Some are intended to limit the damage if a FI goes bankrupt anyway.

Pretty much any kind of business has risks and can go bankrupt. So why are FIs, specifically, subject to these special regulations? Mainly, it is because we believe the bankruptcy of a FI can impose *negative externalities* on other FIs, other businesses and even households. You should remember the concept of externalities and how they can justify regulation from Econ 160 or 360.

In the U.S., banks have been heavily regulated pretty much from the time banking began in America, in the late eighteenth century. In response to the financial crisis of 2008, a federal law, the Dodd-Frank Act of 2010, extended the types of regulation that had long been applied to banks to large nonbank FIs. It also created new types of regulation.

2) How FIs would operate in the absence of regulation

Just about any business faces a tradeoff between risk and profit: most actions that might increase profit will also increase the risk of bankruptcy. The owners of a business must decide how much risk they are willing to take and operate accordingly.

In the specific case of a FI, owners of an FI can earn higher profit by holding less capital, or by holding assets that are longer-term, or less liquid, or more subject to default risk. But all of these actions increase risk of bankruptcy for the reasons described above.

In deciding how much risk to take, the owners of a FI must think not only about how much risk they themselves are willing to take, but also about the preferences of lenders or potential lenders *to* the FI. If lenders or potential lenders think that a FI is taking on too much risk or holding too little capital, they may withdraw their loans or require the FI to pay higher interest. Thus, the amount of risk that a FI takes on may be determined by the "risk preferences" of its owners *and* the people who lend money to the FI.

Possibly, this amount of risk may be optimal (utility or profit-maximizing) from the point of view of the FIs owners and the lenders to the FI. But it may still be *too much risk* from the point of view of society as a whole. That is because the FI's owners and lenders to the FI will not take into account the damage that the FI's bankruptcy might do to *other* businesses and people - the negative externalities.

3) Possible negative externalities of a FI's bankruptcy

a) Contagion

i) Introduction

It has frequently been observed that the occurence of a liquidity crisis at one FI seems to make it more likely that other FIs will suffer liquidity crises. Crisis seems to spread from one FI to others. This is called "contagion." Thus, the occurrence of a liquidity crisis in one FI can set off a wave of liquidity crises and bankruptcy in other FIs. A wave of liquidity crises is called a "financial crisis."

There are at least two reasons such contagion can occur.

ii) Networks of lending among FIs

FIs lend to each other, a lot. Say there are three finaicnal intermediaries, called A, B and C. FI A has been lending to FI B overnight. There is a financial crisis at FI A. So A, desperate for cash, withdraws its overnight loans from B. That can start a financial crisis at B. Suppose C has also been lending overnight to B. When there is a financial crisis at B, C might lose the money it has lent to B. That could make C insolvent. So people who have been lending to C will withdraw their loans to C, creating a financial crisis at C. And so on.

iii) Sales of illiquid assets

Recall that most people are willing to buy illiquid assets only at a low price, but some people, called "knowledgable buyers," would be willing to pay a higher price (because these people have relatively good information about default risk on the assets).

Sometimes the potential knowledgable buyers of one FI's illiquid assets are the same as the potential knowledgable buyers of another FI's illiquid assets. For example, several firms in an industry may have issued bonds that are illiquid because only a few potential bond buyers are knowledgable about business conditions in that industry. If so, sales of illiquid bonds by a FI can depress potential sale prices of other bonds in the same class. That depresses the net worth of FIs holding those other bonds, and worsens the potential fire sale problem for them.

b) Macroeconomic effects

i) Introduction

A wave of bankruptcy in FI's, that is a financial crisis, can cause a recession, or at least worsen a recession that is already underway, by reducing consumption spending and investment spending. In the history of the United States several recessions were caused or worsened by financial crises, including the "Great Recession" of 2008, the "Great Depression" of 1929-33, and the depressions of 1907, 1893, 1884, 1873, and 1857.

ii) By reducing the supply of loans to businesses and households: credit crunch

In section VII) you learned that some businesses and households can borrow only through loans. Most loans are made by financial intermediaries. An FI that has failed, or is on the edge of failure, does not lend. A widespread reduction in lending by FIs in called a "credit crunch." A credit crunch makes it harder for many firms and households to borrow, and hence reduces their spending.

Also, when a FI goes bankrupt the lenders to the FI lose at least some of what they lent to the bank. That reduces their wealth, which tends to reduce their spending.

iii) Additional effect of financial crises that affect banks: no way to make payments

Payments for most sales of goods and services, and most payments of wages and salaries to employees, are not made with cash: they are made using checkable deposits, through banks. Thus, a financial crisis that causes widespread shutdowns of banks tends to reduce spending simply by making it harder to pay for things and employ people.

4) Regulations applied to FIs

a) Introduction

In most countries, government regulation of banks began about as soon as banks began. In America, banks are regulated by several government agencies: a division of the U.S. Treasury called the Office of the Comptroller of the Currency (O.C.C.), established in 1863; the Federal Reserve system, established in 1914; the Fedeal Deposit Insurance Corporation, established in 1933; and also by state agencies. Since the 2008 financial crisis, many of the regulations long applied to banks have been extended to other types of FIs.

b) Capital requirements

Recall that a FI is less likely to become insolvent if it holds more capital.

Left to itself, a FI will choose to hold the amount of capital that maximizes the expected return to the owners' capital investment. This is the amount of capital that just balances balances the cost of holding more capital (at given interest rates, more capital reduces the rate of return to capital invested by the owners) against the possible benefit (the FI might be able to borrow at lower rates of interest).

Capital requirements are regulations that force a FI to hold more capital than it would choose on its own. They specify amounts of capital the FIs must hold in order to be allowed to operate. If a FI's capital falls below the required level, the FI must recapitalize or shut down.

c) Liquidity requirements

A FI that borrows short-term to finance holdings of illiquid assets should not hold illiquid assets *only*. At least some of its assets should be perfectly liquid assets such as Treasury bonds. Liquid assets pay lower returns than illiquid assets. But holding some liquid assets will allow the FI to survive a small run. Up to a point the FI can cover withdrawn loans with sales of its liquid assets; it won't have to sell off illiquid assets in fire sales. Perfectly liquid assets held by a FI are called "secondary reserves."

Liquidity requirements are regulations that force a FI to hold a larger share of its assets in liquid assets (secondary reserves) than it would choose to hold on its own.

In the U.S., liquidity requirements were first introduced by the Dodd-Frank Act.

d) Stress tests

In finance, a "stress test" is an exercise in which people estimate how a specified macroeconomic scenario would affect a FI's net worth, given the FI's current assets and liabilities. In response to results of a stress test, regulators can require a bank to take actions to increase capital or change its asset allocation to reduce risk. The Dodd-Frank Act of 2010 introduced a requirement that large American banks have been required to conduct stress tests twice a year under supervision of Federal Reserve staff, using scenarios made up by the Federal Reserve. Here is one of the scenarios the Federal Reserve posed to large banks in 2019:

... a severe global recession accompanied by a period of heightened stress in commercial real estate markets and corporate debt markets. This is a hypothetical scenario designed to assess the strength of banking organizations and their resilience to unfavorable economic conditions and does not represent a forecast of the Federal Reserve.

The U.S. unemployment rate climbs to a peak of 10 percent in the third quarter of 2020...real GDP falls about 8 percent from its pre-recession peak, reaching a trough in the third quarter of 2020. The decline in activity is accompanied by a lower headline CPI inflation, which falls to about 1¼ per-cent at an annual rate in the first quarter of 2019 and then rises gradually to about 2 percent at an annual rate by the second half of 2020....the interest rate for 3-month Treasury bills falls 2¼ percentage points and remains near zero through the end of the scenario. The 10-year Treasury yield falls by a somewhat smaller amount, resulting in a mildly steeper yield curve. The 10-year Treasury yield reaches a trough of about ¾ percent in the first quarter of 2019 and rises gradually thereafter to 1½ percent by the first quarter of 2021 and 1¾ percent by the first quarter of 2022. Financial conditions in corporate and real estate lending markets are stressed severely. The spread between yields on investment-grade corporate bonds and yields on long-term Treasury securities widens to 5½ percent by the third quarter of 2019, an increase of 3½ percentage points relative to the fourth quarter of 2018. The spread between mortgage rates and 10-year Treasury yields widens to 3½ percentage points over the same time period....Equity prices fall 50 percent through the end of 2019....House prices and commercial real estate prices also experience large declines of about 25 percent and 35 percent, respectively....The international component of this scenario features severe recessions in the euro area, United Kingdom, and Japan, and a shallow recession in developing Asia....The U.S. dollar appreciates against the euro, the pound sterling, and the currencies of developing Asia, but depreciates modestly against the yen because of flight-to-safety capital flows.

e) Deposit insurance

Deposit insurance is just what it sounds like. A bank is required to make regular periodic payments into a fund managed by a government agency. If a bank fails, the insurance fund is used to pay the bank's depositors what the banks owes them. If deposits are fully insured, no depositor need worry that a bank will be unable to pay off on his deposit. Thus, there is no reason to run on the bank.

In the United States, the Roosevelt administration established a federally-controlled company in 1933, the Federal Deposit Insurance Corporation (FDIC), to insure deposits in banks. A bank makes regular insurance payments to the FDIC that increase with the amount of a bank's assets. The FDIC, assisted by other agencies including the Fed, monitors the condition of banks. If the FDIC suspects that a bank is in trouble it immediately closes down the bank and examines its books. If it concludes that the bank cannot continue to operate the FDIC arranges the bank's demise as described in the next section. It pays off the depositors up to a maximum of $250,000. If a deposit is greater than $250,000, the depositor is paid back just $250,000 and loses the rest. Meanwhile the FDIC sells the bank's assets. As much as possible of the FDIC's payments to depositors will be covered by sales of the bank's assets. The rest is covered by the insurance fund. Sometimes a closed bank's assets are more than enough to cover

its deposits. In that case the FDIC pays some money to the people who loaned money to the bank through loans or by buying bonds issued by the bank. But the bank's owners get nothing.

Economists are of two minds with respect to deposit insurance. On the one hand, deposit insurance tends to prevent financial crises from affecting banks and thus lessens the macroeconomic damage of financial crises. On the other hand, deposit insurance encourages banks to take on more risk, because it eliminates a depositors' incentive to monitor the amount of risk taken by a bank and withdraw his deposit if the bank takes too much risk. There is a tradeoff here. That is one reason for the $250,000 limit on FDIC insurance: it preserves the incentive of people with very big deposits to monitor banks' risk.

f) Special bankruptcy procedures for FIs: "resolution authority"

Ordinary bankruptcy, overseen by ordinary bankruptcy courts, is a slow procedure and an uncertain one for creditors. A creditor might get most of the money owed him or very little, depending on the proceeds of liquidating the bankrupt's assets and the court's decisions as to which creditors have first claim on the funds.

Slow and uncertain resolution of a FI's bankruptcy poses special problems. It worsens the problem of contagion. In the end some lenders to the FI will get most of their money back. But until the court decides who gets what, any FI that had been making loans to the bankrupt FI appears to be in danger of taking a bath. A FI that loses the money it lent to the bankrupt FI is that much closer to insolvency itself. Even if only *some* of the FIs who lent to the bankrupt FI will have to take a bath in the end, there may be runs on *all* of them, because people don't know exactly who will have to take a bath.

Better to settle things quickly. If everyone knows quickly exactly which of the FIs that had been lending to the bankrupt FI will be hurt, there may be runs on the FIs that are hurt. But there will not be runs on the FIs that get most of their money back.

For this reason, since the creation of the FDIC, the FDIC, not courts, has handled the bankruptcy of banks. This function of the FDIC is called "resolution authority." The Dodd-Frank Act of 2010 extended the FDIC's resolution authority to cover some large nonbank FIs.

Usually, the FDIC determines the liabilities of a failed FI that will *not* be paid, leaving the failed FI's balance sheet with a small positive value. Then the FDIC tries to sell the remaining liabilities and assets as a package to another FI (rather than trying to sell off the failed FI's assets one by one). As part of this process it can be necessary for the FDIC or Fed to make a loan to the purchaser of the failed FI, collateralized by the failed FI's assets. It can also be necessary for the FDIC to guarantee that it will cover any future losses to the purchaser of the failed FI due to future defaults on the failed FI's assets. Such a guarantee is called a "Shared Loss Agreement."

To make it easier for the FDIC to handle a FI's bankruptcy, large banks and some large nonbank FIs are required to create bankruptcy plans *in advance* and periodically submit the plans to the Fed and the FDIC. Informally, these plans are called "living wills." In the words of the Federal Reserve Board,

Each plan... must describe the company's strategy for rapid and orderly resolution in the event of material financial distress or failure of the company, as well as include both public and confidential sections....Currently, the largest, most complex banking organizations are required to file a resolution plan every other year. Other large domestic and foreign banking organizations are required to file a resolution plan every three years. A third group of firms are required to submit abbreviated resolution plans every three years.

The FDIC and Fed can use a FI's living will to help determine who gets paid what if the FI fails.

©**Christopher Hanes   Economics 450   Class notes**

X) Central banks: introduction
      A) Examples of central banks
      B) More about the Federal Reserve system
      C) Central bank operations
            1) Central bank "tools"
            2) Loans
            3) Open-market operations
            4) Reserve accounts
      D) What central banks do with these tools
            1) Act as "lender of last resort" to stop liquidity crises
            2) Control interest rates
      E) What is a central bank trying to achieve by controlling interest rates?
            1) Before the 1930s: fix exchange rates through the gold standard
            2) Late 1940s-1973: fix exchange rates through the Bretton Woods system
            3) Nowadays: "inflation targeting"
      F) What we will do in the rest of the course

X) Central banks: introduction

A) Examples of central banks

Most countries nowadays have an institution called a "central bank." Many central banks were originally set up as privately owned corporations, granted special privileges in exchange for an extra degree of government control. Now central banks are "owned" by their governments. Examples of central banks are:

The Federal Reserve system or "Fed."

The Bank of England (for the United Kingdom)

The Bank of Japan

The Reserve Bank of Canada

The Reserve Bank of New Zealand

The Reserve Bank of Australia

The European Central Bank (for all the countries that use the euro)

A central bank looks like an ordinary banks in that it takes deposits, buys bonds and makes loans. It makes a profit most of the time. (Their profits go the the national treasury, like taxes.) But a central bank does not aim to maximize its profit. It aims at other things.

B) More about the Federal Reserve system

The Fed is actually twelve separate "Federal Reserve Banks" located in, and named for, twelve American cities that were important centers of banking back in 1912 when the system was planned. The cities are Boston, New York, Philadelphia, Richmond, Atlanta, Cleveland, Chicago, St. Louis, Kansas City, Minneapolis, Dallas and San Francisco. The head of a Federal Reserve bank is called its "president."

The twelve Federal Reserve banks are overseen by a committee that meets in Washington, D.C. Members of the committee are called "governors." They are appointed by the president, confirmed by the senate. The committee is called the "Board of Governors of the Federal Reserve system." The Board of Governors is headed by a "chair" (that is, a chairman or chairwoman).

Most of the Fed's important decisions are made by another committee called the Federal Open Market Committee (FOMC). The FOMC includes all the members of the Board Of Governors plus the president of the Federal Reserve Bank of New York and a subset of the other eleven presidents who rotate through. The chair of the Board of Governors is also the chair of the FOMC.

Why are there twelve separate banks plus a board, rather than just one central bank? No good reason; just 1912 politics.

C) Central bank operations

1) Central bank "tools"

The things that central banks do from day to day, the actual operations they can engage in, are called the "tools" of a central bank

2) Loans

One of a central bank's tools is loans. Nearly all central banks make loans to banks. Many make loans to their governments too (the Federal Reserve does not).

3) Open-market operations

In countries with active bond markets - that includes the United States, of course - central banks buy and sell bonds issued by the government and domestic corporations. This is called "open market operations." Some central banks (not including the Fed) regularly buy and sell foreign bonds too.

4) Reserve accounts

Central banks take deposits from governments and banks (usually not from other types of businesses or households).

In the United States, the accounts that banks hold in the Federal Reserve system are called "reserve accounts." Banks use reserve accounts to get cash for their customers. Suppose, for example, that a bank needs $1,000,000 to replenish its ATMs. The bank will withdraw $1,000,000 from its reserve account and send a truck to its local Federal Reserve bank to pick up the cash.

Banks also use their reserve accounts to make payments to/receive payments from other banks and the Federal government. The Federal government uses its account at the Fed to make and receive payments, too. For example, when your parents pay income tax with a check or electronic payment, their bank makes the associated payment to the Federal government by transferring dollars from the bank's reserve account to the government's Fed account.

D) What central banks do with these tools

1) Act as "lender of last resort" to stop financial crises

Central banks act to prevent or at least ameliorate financial crises. A central bank does this by giving money to financial intermediaries suffering from, or in danger of, a liquidity crisis, in exchange for the FI's illiquid assets. To make this exchange, a central bank simply buys the FI's illiquid assets or (more often) takes the illiquid assets as collateral for loans to the FI. This is called "acting as a lender of last resort" (even if the central banks is simply buying the assets rather than taking them as collateral for a loan).

2) Control interest rates

Recall that all interest rates and yields in an economy are affected by current and expected future overnight rates and term premiums. For example, the yield on a zero-coupon corporate bond of maturity m is:

$$_m i_t = E\left[\frac{_0 i_d + _0 i_{d+1} + _0 i_{d+2} + \dots\dots + _0 i_{d+\dots}}{m \bullet 365}\right] + {_m\tau_t} + default\ risk\ premium + liquidity\ premium$$

The main way that most central banks attempt to control the general level of interest rates is by exerting very tight control over the overnight rate. A central bank controls the overnight rate by choosing a "target" value for the overnight rate and manipulating its tools - loans, open-market operations, rules applied to reserve accounts - to hold the market overnight rate at that target. They are very good at this. From day to day the market overnight rate may deviate a bit from the central bank's target overnight rate, but it will hit the central bank's target on average. People know this is true. So the public's expected value for what the overnight rate will be at a future point in time is just equal to the public's expected value for what the central bank's target overnight rate will be at that point in time. Let i^T denote the central bank's target overnight rate. Then we can say:

$$_m i_t = E\left[\frac{i^T_d + i^T_{d+1} + i^T_{d+2} + \dots\dots}{m \bullet 365}\right] + {_m\tau_t} + default\ risk\ premium + liquidity\ premium$$

Note that what matters is not so much today's target overnight rate, what matters is mainly the public's expectations of what the target overnight rate will be in the future. Thus, central banks try to influence the public's expectations of what the central bank will do in the future. This is called "forward guidance."

Sometimes central banks also try to influence term premiums. One type of operation used to influence term premiums is called "quantitative easing."

E) What is a central bank trying to achieve by controlling interest rates?

1) Before the 1930s: fix exchange rates through the gold standard

Before the First World War, the aim of most central banks was to maintain a commodity-money system called the "international gold standard." In this system a central bank would freely exchange a unit of its country's currency - cash or reserve account at the central bank - for a specified, unvarying amount of gold. The Bank of England, for example, would give four pounds and five shillings (4.25 pounds) for one troy ounce of gold, and sell gold for the same price (the Bank would give you one troy ounce of gold if you gave the Bank 4.25 pounds). The Federal reserve system was originally set up to do the same thing. A Federal Reserve bank would give $20.67 for one troy ounce of gold, or sell gold at the same price. The ratio of country A's gold price to country B's gold price determined the rate of foreign exchange between counry A's currency and country B's currency. Thus one British pound cost $4.87; one dollar cost .205 pounds (about four shillings). Paper money and reserve accounts (central-bank ledger money) was not fiat money, it was token money that could always be exchanged for gold coins or bullion.

To maintain its ability to exchange gold for currency, a central bank had to keep a large supply of gold on hand. It couldn't do that if there was a persistent "drain" of gold, that is if people persistently wanted to buy gold from the central bank. A central bank used its tools mainly to prevent this from happening. To stop a drain of gold, a central bank raised the country's interest rates. In the opposite case, if people persistently wanted to *sell* gold to the central bank, a central bank was free to lower the country's interest rates.

During the First World War central banks stopped exchanging gold for currency. They resumed after the war, but the system broke down permanently in the 1930s.

2) Late 1940s-1973: fix exchange rates through the Bretton Woods system

After the Second World War, until 1973, the aim of most central banks was to maintain a fixed rate of foreign exchange against the U.S. dollar, while the U.S. Treasury was supposed to exchange dollars for gold as under the gold standard. This was called the "Bretton Woods system" because it was set up at an international conference held in Bretton Woods, New Hampshire. Under the Bretton Woods system, the Fed's responsibilities were different from the responsibilities of foriegn central banks.

A foreign central bank was supposed to maintain its ability to exchange its currency for dollars. To do that it had to keep a large supply of dollars on hand. It couldn't do that if there was a persistent drain of dollars from the foreign central bank, that is if people persistently wanted to buy dollars from the central bank. A foreign central bank used adjustments in interest rates mainly to prevent this from happening.

Meanwhile, the Fed was supposed to prevent a drain of gold as under the gold standard. Over the 1960s and early 1970s the Federal Reserve failed to do this. In 1971, in the face of a persistent gold drain, President Nixon ordered the Treasury to stop exchanging gold for dollars. That was the end of the Bretton Woods system.

3) Nowadays: "inflation targeting"

Nowadays central banks of important countries don't try to fix exchange rates in any way; they let markets determine exchange rates (this is called "floating" exchange rates). A central bank's goal is to keep the country's annual rate of inflation in the country at a target value and, subject to that inflation target, keep unemployment at low as possible. In most cases the inflation target is a number about equal to 2 percent. In the United States, the Federal Reserve aims to keep the annual rate of inflation in the PCE price index equal to 2 percent.

Central banks believe that they can influence the rate of inflation in an economy by controlling the general level of interest rates that businesses and households must pay for their borrowing - yields on bonds, interest rates on loans. Macroeconomic theory and historical experience suggest that an increase in interest rates tends to raise unemployment and reduce inflation; a decrease in interest rates has the opposite effect.

F) What we will do in the rest of the course

The rest of the course will be about central banks, especially the Fed. First we will describe what a central bank does as a lender of last resort. Then we will describe how a central bank controls interest rates. Finally we will describe inflation targeting: how the level of interest rates is related to unemployment and inflation and how the Fed responds to events to keep inflation on target.

©**Christopher Hanes Economics 450   Class notes**

XI) The central bank as lender of last resort (LOLR)
    A) Review and introduction
        1) Review
        2) Central banks want to stop liquidity crises from occurring
        3) What about that interest-rate risk stuff?
    B) How to be a LOLR
        1) A run-on FI needs cash
        2) Buy illiquid assets from FI in trouble
        3) Make loans to FI in trouble, taking illiquid assets as collateral
        4) Why does a LOLR have to be a central bank?
        5) The presence of a LOLR can stop liquidity crises from happening in the first place
    C) LOLR best practice: Bagehot's rules
        1) Introduction
        2) Distinguish between illiquidity and ordinary insolvency
        3) Charge a penalty rate for LOLR loans
    D) A problem in LOLR lending: stigma
        1) Definition
        2) Something that makes the problem worse: publicizing who borrows
    E) Other central bank policies related to financial crises (not LOLR but related)
        1) Operations to maintain liquidity of bonds
        2) Lending to non-FI businesses
        3) Reviving banks and other FIs after a big crisis
            a) Aftermath of a big crisis
            b) Solvency testing and certification
            c) Recapitalization

XI) The central bank as lender of last resort (LOLR)

A) Review and introduction

1) Review

Here are some things you need to remember from section IX).

A liquidity crisis can occur when a financial intermediary (FI) has been borrowing short-term (e.g. taking deposits) to acquire illiquid assets (e.g. make loans). The value of illiquid assets is high if you can hold them while they pay off, or sell them off slowly. Their value is low if you must sell them fast in a fire sale. A liquidity crisis is set off when the short-term lenders to the FI withdraw their short-term loans (e.g. when depositors in a bank withdraw their deposits) - when they "run."

If the FI does not have enough liquid assets to cover the withdrawals, it must fire sale its illiquid assets, which makes it insolvent; it can't get enough from fast sale of the illiquid assets to pay off all the lenders to the FI

A liquidity cisis in one FI can spread to others through "contagion," as an FI in trouble withdraws short-term loans it made to other FIs. Also, an FI's sales of illiquid assets in a fire sale depresses their prices, endangering the solvency of other FI's holding the same types of assets.

A wave of liquidity crises is a financial crisis. In a financial crisis many firms that could ordinarily issue bonds find that they cannot do so, while it becomes much harder for firms and households that cannot issue bonds to get loans from FIs. Thus, there is less investment and consumption spending. A recession can result.

2) Central banks want to stop liquidity crises from occurring

A central bank wants to stop liquidity crises from occurring: liquidity crises at a few FIs can become a financial crisis through contagion and a financial crisis can cause a recession.

Fortunately, a central bank *can* prevent financial crises from occurring, or at least prevent contagion from taking place, or at least prevent a financial crisis from becoming really big, by acting as "lender of last resort" (LOLR). A LOLR breaks one of the links in the chain of events that is a liquidity crisis.

3) What about that interest-rate risk stuff?

Back in section IX) I told you about a problem a FI could get into resulting from a combination of interest-rate risk and a run. In many ways that problem was similar to a liquidity crisis. You might think that a central bank might want to act as a lender of last resort to help a FI that is insolvent due to an increase in long-term bond yields, but would still able to pay the interest on its deposits if a run could be stopped.

In fact, an interest-rate-risk-plus-run crisis occurred in the U.S. in March 2023, in a bank called the Silicon Valley Bank, and the Fed took actions related to its functions as lender of last resort. But Federal Reserve officials have not talked much about this, and it is hard to find similar cases in earlier history. By contrast, Federal Reserve officials and economists have talked and written a great deal about liquidity crises and a lender of last resort. So in these notes I will only discuss liquidity crises.

B) How to be a LOLR

1) A run-on FI needs cash

When there is a mass withdrawal of short-term lending to a FI, the FI has to get its hands on a lot of money fast to pay off the lenders. A LOLR does this by giving money to a FI suffering from, or in danger of, a liquidity crisis, in exchange for the FI's illiquid assets. The LOLR can do this in two ways.

2) Buy the FI's illiquid assets at non-fire sale prices

One thing the LOLR can do is buy illiquid assets from a FI suffering a run, not at low fire-sale prices, but at the higher prices one can get if one has lots of time to sell the assets. The FI then has enough money to pay off lenders to the FI; it is not insolvent.

Does the LOLR lose money on this operation? No. Remember, the LOLR is not "overpaying" for the assets. It is paying the same prices a knowledgeable buyer would pay under ordinary circumstances (that is, in the absence of a financial crisis). The LOLR can simply hold on to the illiquid assets it bought from the FI, getting the payments that the borrowers make over time. Or it can sell the illiquid assets off slowly, getting high non-fire sale prices for them.

3) Make loans to the FI taking its illiquid assets as collateral

This also gives the FI cash to pay off lenders to FI, but the FI keeps ownership of the illiquid assets. This is good especially if the illiquid assets we're talking about are loans. The holder of a loan must do monitoring and perhaps invoke restrictive covenants. A central bank doesn't want to mess with that.

To give the FI money without taking ownership of the illiquid assets, an LOLR can *lend* money to the FI taking the illliquid assets as *collateral* for the loan. The amount of money the LOLR lends the FI is equal to the value of the assets at high non-fire sale prices, minus a small deduction to allow for the possibility that the non-fire sale prices of the assets might fall in the future. Thus, there is no way that the LOLR can lose money on the loan. If the FI goes bankrupt despite getting a loan from the LOLR, the LOLR can just seize the collateral (and then hold on to the illiquid assets for the income they generate or sell the illiquid assets off slowly).

Doing 2) and 3) are both called "acting as lender of last resort." Why isn't it called "acting as illiquid-asset-buyer-of-last-resort? Because LOLRs *usually* do 3).

4) Why does a LOLR have to be a central bank?

As I have explained, a LOLR usually makes money on the operation. So why is it up to the central bank to act as LOLR? Why can't privately owned FIs, such as big banks, act as LOLR? In fact, privately-owned FIs *have* acted as LOLR at times in the past. In 1907, before the creation of the Fed, a consortium of New York banks organized by J.P. Morgan acted as LOLR to stop a wave of liquidity crises (the "Panic of 1907").

But a LOLR needs to be able to provide a lot of cash on short notice. A privately-owned FI or even a consortium of FIs wouldn't have enough cash on hand to stop a big wave of crises.

A central bank, on the other hand, creates money by making loans or buying things. So it can always act as LOLR, on whatever scale is needed.

Note that the loans or asset purchases that a central bank makes, acting as LOLR, increase reserve supply.

5) The presence of a LOLR can stop liquidity crises from happening in the first place

If people know that there is a central bank willing and able to act as LOLR, they will not be afraid that a FI will be unable to repay short-term loans to the FI. Runs won't happen.

C) LOLR best practice: Bagehot's rules

1) Introduction

Walter Bagehot, pronounced badge-it (rhymes with gadget), was the editor of a British business/finance magazine (*the Economist*, still around today), sort of the British *Wall Street Journal*, in the middle of the 19th century. He wrote about what the Bank of England did in liquidity crises. What the Bank was doing was acting as lender of last resort. Bagehot said: "what the Bank ought to do is this..." His recommendations are now called "Bagehot's rules" for being a good LOLR.

2) Distinguish between illiquidity and ordinary insolvency

An FI is in trouble. Do you, the central bank, help it by buying its illiquid assets or taking the illiquid assets as collateral for a loan to the FI?

Bagehot said, *do not* help the FI if its problem is ordinary insolvency due to defaults on its assets or interest-rate risk. Such FI's should be left to die, so that FIs have incentive to not acquire too-risky assets.

Do help the FI if it would be solvent except for the liquidity crisis/fire sale problem.

3) Charge a penalty rate for LOLR loans

"Penalty rate" means higher-than-normal interest rate. You want the liquidity crisis to cause the FI some pain, but not to kill it. This gives FI's incentive to avoid liquidity crises as much as possible.

D) A problem in LOLR lending: stigma

1) Definition

Central banks want an FI suffering a run to come for LOLR help early, before many of the lenders to the FI have run. "Stigma" can prevent an FI from taking LOLR help.

Stigma is a word that generally means "a mark of shame." (Its original meaning was one of the nail holes in Jesus' hands; plural stigmata. Greek.)

In this context stigma means: if people see a FI taking a loan from a central bank, they will fear the FI is on the edge of illiquidity/insolvency and the run on the FI will intensify, making things worse.

FI's fear this will happen. So FI's might not come to the central bank for help early, when a run is just starting; they wait until the last minute, or when it is too late to save the FI.

2) Something that makes the problem worse: publicizing who borrows

You need to keep it secret who has borrowed from the central bank, if you can. Believe it or not, Congress has twice passed laws requiring the Fed to PUBLISH the identities of borrowers.

F) Other central bank operations related in financial crises (not LOLR but related)

1) Operations to maintain liquidity of bonds

Recall that a bond can be liquid or illiquid, depending on the number of people willing to buy the bond at about the same price.

There can be multiple equilibria and self-fulfilling expectations here. If everyone thinks the bond will be liquid in the future, everyone is willing to buy that type of bond. It's liquid. If people think the bond may become illiquid, they won't buy it. And it's illiquid!

A central bank can push things toward the good outcome (keeping that class of bond liquid) by promising to buy the bond from anyone who wants to sell it to the central bank, at a non-fire-sale price, or taking that type of bond as collateral for loans, valuing the bond at a non-fire sale price. If everyone knows the central bank will do this, more people are willing to buy the bond themselves. And the bond remains liquid.

2) Lending to non-FI businesses

Recall that a financial crisis can shift the IS back by making it hard (even impossible) for many people and businesses to borrow: FIs that would ordinary make loans are out of operation; people are unwilling to buy bonds issued by companies that would ordinarily be able to sell bonds.

Despite the existence of the Fed, big financial crises occurred, in 1933, 2008 and (briefly) 2020. In all of these crises the Fed made direct loans to ordinary companies and/or bought bonds directly from private companies issuing them, so that the companies would be able to borrow and keep functioning.

3) Reviving FIs after a financial crisis

a) Aftermath of a financial crisis

We want people to start lending to FIs again (e.g. deposit in banks), so that the FIs can make loans to businesses that are unable to sell bonds.

But in the wake of a widespread crisis, some FIs will be insolvent; some will be barely solvent; some will be very solvent.

Potential lenders to FIs (e.g., people who had withdrawn their deposits from banks but could now put money back into banks) don't know which FI's are in which category, and thus are unwilling to lend to any FI's. If no one lends to FIs, FIs can lend to companies and people who need to borrow through loans.

To solve this problem, government authorities, usually the central bank with the assistance of the government itself, can engage in "solvency testing and certification" of surviving FIs and "recapitalize" ones that are low on capital. The Fed together with the U.S. Treasury did these things after the 1933 and 2008 financial crises.

b) Solvency testing and certification

This is when the central bank or other government authority sends agents to inspect FIs' accounts to figure out which ones are insolvent, barely solvent (zero or not much capital) or very solvent (lots of capital).

The authority then closes down the insolvent ones, reopens the very solvent ones, and recapitalize the barely-solvent ones.

If people trust the government authority, they will believe that the reopened FIs, certified as solvent by the government authority, really is solvent and well-capitalized, and they will lend to the reopened FIs.

c) Recapitalization

This is when the government authority forces barely-solvent FIs to sell shares to new part-owners. The funds the FI raises by selling shares are used to buy assets. The FI's capital increases (look back at your notes on "capital"). The owners of the FI might be reluctant to raise capital this way, because selling shares - raising capital - reduces existing owners' shares of profits. So sometimes a FI must be *forced* to do this.

Sometimes a government authority wants to save and recapitalize a FI, but private investors are unwilling to buy new shares in the FI. In such cases government authorities have themselves bought shares in FIs to boost the FIs capital. Then the government authority holds on to the shares, taking its share of the profits, until private investors are willing to buy the shares from the government authority.

Christopher Hanes Economics 450 Class notes

XII) Financial Crises and Fed actions in recent years
    A) Introduction
    B) Limits on what the Fed can do
        1) Introduction
        2) The Federal Reserve Act
        3) Political considerations
    C) The financial crisis of 2008
        1) Introduction
        2) Before the crisis: the way things were in the early 1990s
            a) Mortgages, Fannie Mae and Freddie Mac (GSEs)
                i) Mortgages and mortgage-backed securities (MBS)
                ii) Fannie Mae and Freddie Mac
            b) Overnight loans: fed funds and repos
            c) Banks
            d) Nonbank financial companies
                 i) Introduction
                ii) Investment banks
                iii) Money-market mutual funds
        3) The run-up to the crisis
            a) House prices rise
            b) Subprime (nonconforming) mortgages
            c) Mortgage-backed securities made from subprime mortgages
                i) Introduction
                ii) Credit-default swaps
                iii) Subprime mortgage MBS were liquid
            d) Investment banks sart engaging in financial intermediation
            e) Asset-backed commerical paper
            f) Money-market mutual funds
        4) Crisis
            a) House prices fall
            b) Defaults on subprime mortgages
            c) Defaults on MBS made from subprime mortgages
            d) Subprime MBS become illiquid
            e) Fannie Mae and Freddie Mac taken over by the federal government
            f) Runs on financial intermediaries
                i) Banks
                ii) Investment banks
                iii) Money markey mutual funds
        5) The Fed's response
            a) Introduction
            b) Fed loans to banks
            c) Fed loans to investment banks
                i) Lending to investment banks through non-recourse loans to banks
                ii) Bear Stearns and invocation of 13(3)
                iii) Creation of new facilities to lend to investment banks
                iv) Lehman Brothers
                v) Investment banks become banks
            d) AIG
            e) Money market mutual funds

      f) Lending to issuers of commercial paper
      g) Swap lines
      h) The Term Asset-backed Securities Loan Facility (TALF)
    6) Extension of deposit insurance
    7) Reviving FIs: stress tests and recapitalization
  D) The Covid crisis
    1) Introduction
    2) New constriant on Fed lending to nonbanks
    3) The crisis
      a) Similar to 2008
      b) Different from 2008
        i) Withdrawal of "term lending"
        ii) Need to borrow by non-FI businesses and local governments
        iii) Treasury and agencyMBS markets become less liquid
    4) What the Fed did
      a) Restart programs from 2008
      b) To help the Treasury bond market
        i) Open market operations in Treasuries
        ii) Foreign and International Monetary Authorities Repo Facility
      c) Lending to non-FI firms and governments
  E) The March 2023 crisis
    1) Introduction
    2) The crisis
    3) What the Fed did: the Bank Term Funding Program (BTFP)
    4) What the FDIC did: pay off uninsured deposits

XIV) Financial Crises and Fed actions in recent years

A) Introduction

The Federal Reserve has been faced with financial crises several times since it began operating in 1914. The first big one occurred during the Great Depression of 1929-33. This was a wave of liquidity crises in banks. In that crisis the Federal Reserve failed to act as lender of last resort. This failure on the part of the Fed is what made the Great Depression great. To stop the crisis President Roosevelt shut down all the country's banks in an action cheerfully called a "bank holiday." While the banks were closed, regulators examined the books of just about all banks in America. Banks that were obviously solvent were allowed to reopen quickly; obviously insolvent ones were closed; banks that were somewhere in between were required to recapitalize. In this course we do not have time to examine the financial crisis of the Great Depression. But we will examine three crises and near-crises of recent years.

One is the financial crisis of 2008. (This may not seem recent to you, but it does to me.) In this crisis the Fed did act as LOLR and responded in several other ways partly because Fed policymakers were well aware that the Fed had failed after 1929. Unlike the crisis of the Great Depression, the 2008 crisis was primarily a matter of liquidity crises in *nonbank* FIs.

When Covid hit in March 2020, an incipient crisis was headed off by fast Fed actions.

Finally, there was an incipient crisis in spring 2023, with the failures of the Silicon Valley Bank and Signature Bank. This was snuffed out quickly by Fed action. But it is still interesting, because it was

fundamentally unlike the crises of 1933 and 2008. It was not a liquidity crisis. It was instead due to interest-rate risk.

Unfortunately, to tell the story of the crises, I have to mention a number of specific dates, names of specific programs of the Federal Reserve and U.S. Treasury, names of banks, peculiar financial-market jargon, and things like that. You do *not* have to memorize any of this stuff. But to understand the story you will have to keep track of it while you are doing the assigned readings. To help you do this, there is a **Glossary** of terms in the coursepack.

B) Limits on what the Fed can do

1) Introduction

In order to understand what the Fed did in these crises, you need to keep in mind that there are legal and customary limits on what the Fed can do. Some of these limits are part of the Federeal Reserve Act of 1912, the law that created the Fed, and subsequent amendments to the law that Congress and the president have made from time to time. Others are due to political considerations. The Federal Reserve system is not written into the U.S. Constitution. It was created by an ordinary law, and could be destroyed by another law. From the beginning of the Fed to today there have been lots of voters and Congressmen who have been suspicious of the Fed.

2) The Federal Reserve Act

The Federal Reserve Act (in section 10B) requires that a loan made by a Federal Reserve Bank must be "secured to the satisfaction of the Federal Reserve bank." The Fed has generally interpreted that to mean that a loan must be secured by collateral, with a haircut against its market value. Also, debts to the Fed are "senior" to other debts: the borrower must pay them off before it can pay other debts.

The original Federal Reserve Act allowed the Fed to make loans *only* to banks, not to nonbank FIs or non-FI businesses. In 1932 Congress passed a low that, among other things, amended the Federal Reserve Act. It allowed the Fed to lend to nonbanks - nonbank FIs and even ordinary businesses - but only under special circumstances. "In unusual and exigent circumstances, the Federal Reserve Board, by the affirmative vote of not less than five members, may authorize and Federal Reserve Bank" to lend to "any individual, partnership, or corporation..." This amendment is in section 13(3) of the Act.

3) Political considerations

The Fed has avoided making loans or buying assets that would involve too much risk of having to book a big loss on the Fed's balance sheet. This is partly because Federal Reserve officials do not want to appear to have lost "the public's money," and partly because they do not want to appear to be "bailing out" banks or other FI's. They want it to look like their operations actually make money for the public. In fact, most of the time Fed operations do make money. And all profits from Federal Reserve operations, after covering Fed expenses like salaries, go to the U.S. Treasury.

What if the Fed needs to do something to help the economy, but that thing necessarily involves a risk of loss? Then the Fed brings the U.S. Treasury in on the operation and sets up the rules of the operation so that the Treasury, not the Fed, would book most or all of the loss if something very bad happened. The term used for this is "backstop:" the Treasury provides a "backstop" to limit the Fed's possible loss.

Of course, this is actually nonsense, because of what I said before: *any profits earned by Federal Reserve banks go to the U.S. Treasury anyway*. Whether the Treasury takes a loss directly, or indirectly through a

decrease in the profits it gets from the Fed, the Treasury takes the same loss in the end. But most people don't seem to understand that.

C) The financial crisis of 2008

1) Introduction

The immediate cause of the 2008 crisis was an unprecedented natiowide run-up in house prices in the early 2000's, followed by a crash in house prices that began in 2006. I will not attempt to explain this rise and fall of house prices. (Economists are still arguing about this.) The important thing is that the drop in house prices caused a financial crisis. A lot of economists, including those at the Fed, were expecting house prices to crash, but almost no one expected the house price crash to cause a financial crisis. Economists soon figured out that house prices could affect the financial system in this way because of *changes* in financial markets that had occurred starting in the late 1990s, partly in response to the big run-up in house prices of the early 2000s. These changes made nonbank FIs vulnerable to house prices in a new way. Thus the house price crash caused a wave of financial crises in nonbank FIs.

To explain this, I will first describe some aspects of financial markets *before* the late 1990s. Then I will describe the changes that made nonbank FIs vulnrable to house prices. Third, I will describe the events that occurred in response to the drop in house prices that began in 2006, which amounted to a financial crisis. Finally I will describe the Fed's response.

2) Before the crisis: the way things were in the early 1990s

a) Mortgages, Fannie Mae and Freddie Mac (GSEs)

i) Mortgages and mortgage-backed securities (MBS)

A mortgage is a loan that is collateralized by a piece of real estate. If a mortgage borrower defaults on the loan, the lender seizes the piece of real estate in a legal action called "foreclosure." Many people take mortgages to finance the purchase of a house or condo to live in. A "mortgage backed security" (MBS) is a bond made out of mortgages. That is, someone buys a whole lot of mortgages from the original lenders (often banks or credit unions) and at the same time issues a bond. The payments on the mortgage loans are used to pay the coupons and final payment on the bond.

ii) Fannie Mae and Freddie Mac (GSEs)

In the early 1990s there were two corporations, which the federal government had set up in the 1930s, called "Fannie Mae" and "Freddie Mac." (These are nicknames for FNMA and FHLMC, which are acronyms for the companies' real names: the "Federal National Mortgage Association" and the "Federal Home Loan Mortgage Corporation.") Fannie and Freddie were called "government sponsored enterprises (GSEs). They were technically private corporations, owned by shareholders. But no one thought the Federal government would ever let them actually go bankrupt.

The business of Fannie and Freddie was mainly to buy up mortgages, make MBS bonds out of them, and sell the bonds. Fannie and Freddie guaranteed payment on the bonds: that is, they promised to make the payments on a bond even if there were defaults on the mortgages that bond had been made of.

Fannie and Freddie would buy a mortgage only if it was a "conforming" mortgage (also called a "prime" mortgage). A conforming mortgage has certain characteristics, specified by rules set by Fannie and

Freddie, meant to reduce the danger that the borrower will default on the mortgage. The amount lent in a conforming mortgage must not be too high relative to average house prices in the region. It must be substatially less than the full purchase price of the house. That is, if the house cost $500,000, a conforming mortgage loan on that house cannot be more than $450,000. The buyer must cover the rest of the house price with her own money (the "down payment"). A conforming mortgage must be a "fixed-payment" loan (all payments the same size; final payment is not especially big), and the loan must be long-term, typically 30 years. Thus, the monthly payment was predictable and relatively small. Finally, before making a conforming mortgage loan the lender has to check the borrower's monthly income to make sure that the borrower can easily cover the monthly mortgage payment.

In addition to the business of buying up conforming mortgages and making MBS bonds out of them, Fannie and Freddie gotten into another business in the early 2000s. This other business was a kind of financial intermediation. Fannie and Freddie would borrow by issuing bonds. (I am not talking about MBS bonds now; I am talking about ordinary bonds issued by Fannie and Freddie.) Then they would use the funds they got from selling bonds to buy and *hold* some mortgages, some of the MBS bonds they themselves had created, and MBS bonds created by other firms. They could make money on this because the yields on bonds issued by Fannie and Freddie were very low, lower than interest rates on mortgages and yields on MBS bonds. Yields on bonds issued by Fannie and Freddie were so low because people believed the federal govenment would never let Fannie and Freddie go bankrupt.

b) Overnight loans: fed funds and repos (repurchase agreements)

Back in section V) I introduced the concept of an overnight loan. Recall that in an overnight loan, the lender promises to deliver funds by the end of a day and the borrower promises to return the funds with interest on the morning of the following day.

Two types of overnight loans became common in the U.S. after the Second World War.

One is called "federal funds loans," "fed funds loans" for short. These are uncollateralized overnight loans, mainly between banks.

The other is called "overnight repurchase agreements," "overnight repos" for short. These are collateralized in a way that might seem confusing at first. In an overnight repo the lender buys a bond from the borrower at a given price X, promising to deliver the funds by the end of the day. At the same time the borrower and lender sign a contract that says the borrower will buy the bond back from the lender the following morning at a price Y (not tomorrow morning's market price, but the prespecified price Y). If the borrower does not follow through on the promise to buy the bond back, the lender gets to keep the bond. Y is always a little more than X. Do you see how this is really a collateralized overnight loan? The bond is the collateral. X is the amount of the loan. The interest on the loan (expressed on an overnight basis) is $(Y-X)/X$.

A lender could lose money on a loan like this is if the borrower did not follow through on the promise to buy the bond back *and* the market value of the bond on the following morning was lower than the amount of the loan X (in that case, when the lender sold the bond the following morning, he would end up with less than X). To make sure this doesn't happen, overnight repo lenders alway insist that the amount of the loan is a good bit less than market value of collateral (the bond) on the day the loan is made. The difference between market value of the collaetral on the day the loan is made and the amount of the loan X is called a "**haircut.**"

c) Banks

Banks borrowed short-term through deposits and overnight loans. They held illiquid assets (loans and illiquid bonds). Thus they were in danger of liquidity crises and runs. To protect banks from runs, the United States had many regulations and institutions which I have described in earlier sections: capital requirements; deposit insurance (run through the FDIC, the Federal Deposit Insurance Corporation); access to LOLR loans from the Fed; and "resolution authority" (special fast bankruptcy procedures).

d) Nonbank financial companies

i) Introduction

There were many kinds of financial firms other than banks. One important type of nonbank financial firm was called "investment banks." Another type was called "money--market mutual funds." Nonbank financial firms were not subject to the same regulations or have the same protections as banks. But that was OK, because, unlike banks, nonbank financial firms were not potentially vulnerable to liquidity crises and runs. Most nonbank financial firms did not engage in financial intermediation at all - they did not borrow in order to lend. And the nonbank financial firms that *did* engage in financial intermediation held only liquid assets - they were in no danger of the fire sale problem.

ii) Investment banks

In the 1920s, big New York City banks such as J.P. Morgan and Company had a sideline in helping private companies issue bonds and new shares of stock. Sometimes a bank would buy up all (or some) of a company's new stock or bond issue from the company wholesale, then retail the bonds or shares of stocks to investors, taking a markup of course. This business is called "underwriting" bonds. Helping a company design a bond or stock issue and sell the new security, perhaps through underwriting, is called "investment banking." Note that, despite the name, investment banking is *not* a form of banking. It is not about taking deposits and making loans or buying bonds.

In 1933, during the Roosevelt administration, a law (called the "Glass-Steagall Act") decreed that banks could no longer engage in investment banking on the side: a bank could not do investment banking, and a business that did investment banking could not be a bank too. The investment banking arms of big New York banks were split off to become separate corporations called "investment banks." J.P. Morgan's investment banking business became an investment bank called "Morgan Stanley." Some other investment banks were "Goldman Sachs," "Lehman Brothers" and "Bear Stearns." The 1933 law that separated investment banking from banking was repealed in 1999, but as of 2008 most investsment banks were still stand-alone corporations. They were big businesses. Like hedge funds today, they were firms where smart greedy young people wanted to work. (I hope you are smart and greedy, get a job in a hedge fund, and give money to SUNY-Binghamton once you make your fortune.)

In addition to underwriting investment banks advised companies on mergers and acquisitions. Many investment banks had subsidiaries that were bond dealers. The bond-dealing subsidiaries of investment banks were important dealers. Some were "**primary dealers**." A primary dealer is a bond dealer that is licensed to buy and sell from the Fed in Fed open-market operations. (There are bond dealers that are not primary dealers. These cannot buy and sell from the Fed.)

Note that investment banks did not engage in financial intermediation: that is they did not take deposits or borrow short-term to acquire long-term or illiquid assets.

iii) Money-market mutual funds (MMMFs)

A "money-market mutual fund" is a special type of "mutual fund."

In order to understand what a money-market mutual fund does, you first have to understand what is done by an *ordinary* mutual fund, that is a mutual fund that is *not* a "money market mutual fund." An ordinary mutual fund is *not* a financial intermediary. The mutual fund buys stocks, bonds, real estate, or a mix of such assets, and sells shares in the fund. A shareholder in the mutual fund gets quarterly payments equal to her share of the coupon payments on the bonds, or the dividends on the stocks, or the rents on the buildings. If a shareholder sells her share back to the mutual fund company, the shareholder will be paid her share of the current market value of the assets held by the mutual fund. Because the price of the assets might have fallen since she bought her share of the mutual fund, the money that she gets when she sells her shares in the MMMF might be lower than the amount of money she put into the fund originally.

A "money market mutual fund" or MMMF *is* a financial intermediary. Its "shares" are sold to investors at a "price" of $1. The MMMF promises that it will always buy shares back at the same "price" of $1. In the meantime, the MMMF pays the investor a fixed percentage rate of return on each $1 share. In effect, the MMMF is borrowing overnight from the people who buy shares in the fund. The MMMF is making the same promise a bank makes when it takes a deposit: if you put in $1000, you can always take the entire $1000 out again and in the meantime get interest. An investor in a MMMF can also write checks against her "shares" in the mutual fund.

So how is a MMMF *different* from a bank? Unlike a bank, a MMMF does not buy long-term bonds: if it did, and the price of the bonds fell due to a rise in interest rates (interest-rate risk), the MMMF might find itself unable to pay its investors everything it owes them. A MMMFs does not make long-term loans, either. A MMMF *only* makes very short-term loans, mainly overnight loans, and buys bills: Treasury bills and commercial paper (see Glossary). Because all of a MMMFs assets were short-term, interest-rate risk could not be a problem. Because the assets were also liquid, a liquidity crisis could not be a problem - or so it was believed before 2008.

Why are these businesses called "money market" mutual funds? The total market for overnight loans, other very short-term loans (like loans at a maturity of a week or two), and bills is called the "money market." Hence "money market mutual fund."

3) The run-up to the crisis

a) House prices rise

Around 1999-2001 house prices began to rise rapidly. Figure 1 at the end of this section is a graph of "real" house prices, that is an index of house pries divided by an index of prices in general (the CPI). The figure shows that real house prices rose rapidly for several years starting at the end of the 1990s. Previous waves of rising house prices in America had been confined to specific geographic regions. This one was almost nationwide. (The only place where house prices did not rise was upstate New York.)

b) Subprime (nonconforming) mortgages

People believed that house prices would keep rising in the future, for a long time. So many lenders and investors came to believe that the provisions of conforming mortgages were unnecessary to prevent default on the mortgage. If a mortgage borrower couldn't pay off a mortgage out of her income, she could just sell the house at a profit and pay off the mortgage that way!

So lots of lenders began making lots of mortgage loans that were *not* conforming in one way or another. These mortgages were called "subprime" or "nonconforming" mortgages. A nonconforming mortgage loan might have a final payment that was much bigger than the other payments ("balloon loans"). The interest rate on the loan might not be fixed - it might be automatically linked to short-term interest rates, so that it went up if interest rates went up ("floating rate loans"). The monthly payment on the loan might be quite large relative to the borrower's monthly income. Or the term of the term of the loan might be much shorter than thirty years (as short as five years), requiring the borrower to get another mortgage when the first mortgage came due. Paying off an old mortgage with money borrowed in a new mortgage is called "refinancing." Because lenders believed house prices would rise forever, they did not worry that borrowers would default on such mortgages - if worst came to worst, the borrower could just sell the house for a high price and pay off the mortgage.

Of course, nonconforming mortgage loans could not be sold to Fannie and Freddie. And if house prices stopped rising, there would be lots of defaults on subprime mortgages. But nobody worried about that.

c) Mortgage-backed securities made from subprime mortgages

i) Introduction

Financial firms, including investment banks, began to make MBS's out of nonconforming mortgages.

ii) Credit-default swaps (CDS)

Fannie and Freddie guaranteed payment on the MBS bonds they made out of conforming mortgages. Issuers of MBS bonds made out of nonconforming mortgages did not make the same guarantee. But the bonds were frequently sold with a sort of insurance policy attached, called a credit-default swap (CDS). A CDS is a type of "derivative."

Generally, a derivative is a financial contract in which the payments that the parties promise to make in the contract depend on what happens to a bond, loan or other financial contract (hence "derivative" - the derivative contract is *derived* from another deal).

In a CDS, party A promises to pay party B some money if there is a default on a specified loan or bond. If party B is the owner of the bond or loan in question, then party B has effectively purchased an insurance policy against default on the bond or loan, like the fire insurance your parents have on their house on Long Island. Party A is like an insurance company. But in a CDS party A doesn't have to be an insurance company. Many investment banks and other types of financial firms acted as party A in CDS.

Subprime-mortgage MBS's were often sold along with a CDS contract that promised a payment to the holder of the MBS if there was a default on the MBS. The CDS contract was meant to serve as insurance against default on the subprime MBS's, to make them more similar to prime MBSs made by Fannie ansd Freddie, which were guaranteed by those companies.

Of course, there was a risk that the companies who had "insured" the subprime MBS bonds by being "party A" in a CDS contract might themselves default on the deal, that is not pay off as promised in case of default on the MBS. The risk that party A in a CDS fails to provide the insurance as promised is called "counterparty risk." But no one was very worried about this (before 2006).

iii) Subprime-mortgage MBS's were liquid

Remember that everyone thought house prices would go on rising. Even though these MBS's were constructed from subprime mortgages, people did not think there would be lots of defaults on these mortgages, so they thought the probability of default on the bonds was low. And the bonds were "insured" by the CDS anyway.

Because people did not worry about default risk on the bonds, there was no "lemons problem" on the bonds. They were very liquid.

d) Investment banks start engaging in financial intermediation

Recall that I said that some investment banks were making subprime MBS to sell. As part of this process, they held large inventories of MBS bonds they had not yet sold, the way a shoe store holds an inventory of shoes. This was called "warehousing" the bonds. To finance these inventories, investment banks borrowed through overnight repo. As the collateral for this repo borrowing, the investment banks used the MBS bonds themselves. This was normal behavior for an investment bank. Investment banks had always held inventories of bonds they were helping companies to issue.

As the subprime MBS boom went on, however, investment banks began to do something different. They began to borrow overnight through repo loans to fund acquisition of MBS bonds to *hold* as an investment, using the MBS bonds themselves as collateral for the repo loans. This was a form of financial intermediation!

e) Asset-backed commercial paper

Commercial paper is a short-term bond (three months to a year) issued by a private company. Many companies issue commercial paper to pay for inventories of raw materials (to manufacture things) or stocks of goods to sell (retailers).

In the run-up to the crisis, banks and investment banks began to issue commercial paper and using the money raised that way to buy long-term bonds, especially subprime MBS's. Commercial paper issued for this purpose was called "asset backed commercial paper" (the "assets" referred to were the bonds). This was a sort of financial intermediation, of course.

The banks and investment banks involved did this in a tricky way. They did not want to be on the hook to pay off the commerical paper if the bonds went bad. So they set up separate corporations called "shell corporations," controlled by the bank or investment bank but legally a separate corporation. The shell corporation did nothing but issue the commerical paper and buy and hold the bonds. The bank or investment bank that had set up the shell corporation would keep the profits from the spread between the return to the bonds and the yield paid on the commercial paper. But if the bonds went bad and the shell corporation was bankrupt, the bank or investment bank was not liable to pay off the commerical paper.

f) Money-market mutual funds

They bought the asset-backed commercial paper issued by shell corporations to buy subprime MBS's.

4) Crisis

a) House prices fall

See Figure 1 at the end of the section.

b) Defaults on subprime mortgages

Lots of defaults.

c) Defaults on MBS made from subprime mortgages

Of course.

d) Subprime MBS's become illiquid

Subprime MBS's became illiquid as investors realized that there was, in fact, considerable default risk on subprime MBS's, and that someone else might know more about that default risk - lemons problem! Investors also began to realize that the companies who had "insured" the bonds by being "party A" in a CDS contract might themselves default on the deal, that is not pay off as promised in case of default on the MBS. The risk that party A in a CDS fails to provide the insurance as promised is called "counterparty risk."

e) Fannie Mae and Freddie Mac taken over by the Federal government

People began to worry that something would happen to Fannie Mae and Freddie Mac. Starting in the late 1990s, Fannie and Freddie had begun to buy and hold MBS made by other firms out of nonconforming mortgages, to profit from the spread between the yields on these bonds and bonds issued by Fannie and Freddie. Now, the market prices of these bonds were plummeting. Even worse: there were begining to be lots of defaults and foreclosures on *conforming* mortgages that Fannie and Freddie had bought to make into MBS - recall that Fannie and Freddie guaranteed payments on these bonds.

Prices of shares (stock) in Fannie and Freddie plummeted. Fannie and Freddie both had very little capital, so it was not clear how they could stay in business. If they went out of business, the whole system of mortgage lending in the U.S. would cease to function. People would not be able to buy houses. That would make house prices fall even more.

Eventually, in September 2008, federal authorities placed Fannie and Freddie into "conservatorship." Shareholders in the firms lost their investments. Management was fired and replaced. The U.S. Treasury recapitalized the firms by buying new shares of stock in the firms.

As of summer 2024, the two firms are still under government conservatorship.

f) Runs on financial intermediaries

i) Banks

People began to worry that banks holding subprime MBS's would go under. Most deposits in banks were insured by the FDIC so there was not much withdrawal of deposits. But there were large withdrawals of

other types of short-term loans to banks. This caused a "credit crunch": banks pulled back on lending, even to companies they would ordinarily have been happy to lend to.

ii) Investment banks

Recall that investment banks were holding lots of subprime MBS. There were runs on investment banks: people who had been lending to the investment banks overnight refused to roll over their loans. Meanwhile the subprime MBS had become illiquid. So, liquidity crises! And contagion: many financial intermedaries, including banks, had been lending to these investment banks.

Some of the runs on overnight loans to investment banks took the ordinary form: lenders simply withdrew their loans. But others took a peculiar form related to haircuts. Recall that the overnight borrowing through investment banks had been through overnight repos, and repo loans are made with *haircuts*: the amount of a repo loan is always less than the current market value of the bond that is the collateral for the loan. Some lenders offered to keep lending to an investment bank, but demanded bigger haircuts on the loans, like this: "Yesterday, I was willing to lend you 1 million dollars overnight on the collateral of bonds worth 1.25 million. [The .25 million was the haircut.] Today, I will still lend you 1 million dollars. But you will need to give me as collateral bonds worth 2 million." This was reasonable from lenders' point of view: it would be a better safeguard against a possible drop in the market value of the collateral bonds. But it meant that an investment bank, holding a given amount of bonds that could be used as collateral to borrow, could not borrow as much. It was equivalent to a withdrawal of overnight lending to the investment bank.

iii) Money-market mutual funds

As I said earlier, MMMFs had bought the asset-backed commercial paper issued by structured investment vehicles to acquire subprime MBS. As people began to fear defaults on the bonds they began to fear defaults on the asset-backed commercial paper so they began to fear defaults by the MMMFs and they withdrew their deposits from the MMMFs.

5) The Fed's response

a) Introduction

In section XI) I discussed several things that policymakers can do to stop a financial crisis and ameliorate its effects, including:
- act as lender of last resort
- operations to maintain liquidity of bonds
- lend to non-FI businesses
- recapitalize FIs

In response to the crisis of 2008, the Fed did the first three, sometimes with help from the U.S. Treasury. Recapitalization of banks was done by the Treasury alone..

In addition to the actions the Fed took to stop the financial crisis, it took the usual action to fight a recession: engineering a decrease in interest rates. I will not discuss that in this section, which is focused on financial crises. We will return to Fed interest-rate policy in a later section.

b) Fed loans to banks

In response to the withdrawal of short-term lending to banks, the Fed made a lot of loans to banks. Fed loans to banks are often called "discount window loans," for a purely historical reason I will not bore you with.

In section XI) I said that LOLR lending by a central bank can be hindered by stigma (banks might refuse to borrow because they would fear that borrowing from the Fed would just cause more withdrawal of loans from the bank because lenders to the bank would take borrowing from the Fed as a signal that the bankk was in trouble).

To get around this, Fed officials designed a new form of LOLR lending in a program called the **Term Auction Facility (TAF)**. Essentially, the TAF allowed a bank to borrow from the Fed *anonymously*, in a deliberately complicated way, through a sort of auction. Quoting from the Fed website, "All loans extended under the TAF were fully collateralized, and the funds were allocated through an auction, in which participating depository institutions placed bids specifying an amount of funds, up to a pre-specified limit, and an interest rate that they would be willing to pay for such funds. The funds were allocated beginning with the highest interest rate offered until either all funds were allocated or all bids were satisfied. All borrowing institutions paid the same interest rate, either the rate associated with the bid that would fully subscribe the auction, or in the case that total bids were less than the amount of funds offered, the lowest rate that was bid." (I told you it was complicated.)

c) Fed loans to investment banks

i) Lending to investment banks through non-recourse loans to banks

Fed policymakers wanted to make loans to investment banks but they could not do so without invoking section 13(3). They did not want to invoke section 13(3) because that would require them to declare that they were in "unusual and exigent circumstances," and that might cause greater panic in financial markets.

At first the Fed engaged in some pretty fancy operations to make loans to investment banks indirectly, through a bank. To do this, the Fed would identify collateral that the investment bank could use to borrow. The cooperating bank would make a loan to the investment bank on that collatreal. the Fed would simultaneously make a loan to the bank of the same amount, on the same collateral. Of course, the bank would be willing to cooperate in this operation only if the could not possibly lose money on the deal. To satisfy the cooperarting bank in this regard, the Fed made its loan to the bank a "non-recourse loan." That is, the loan specified that if the bank defaulted on the loan the lender (that is, the Fed) could seize the collatereal for the loan (the sutff originally put up by the investment bank) but not any more assets of the borrower (the bank).

ii) Bear Stearns and invocation of 13(3)

One investment bank that was subject to runs in a spectacular way was called "Bear Stearns." Fed policymakers had to decide whether to get involved. In doing so, they thought about Bagehot's rules. They judged that Bear Stearns was fundamentally solvent, just illiquid. They arranged for J.P. Morgan, a bank, to take over Bear Stearns. , by making a loan to J.P.Morgan on the collateral of Bear Stearns' assets. Thanks to this deal and the loan from the Fed, Bear Stearns did not go bankrupt; the people who

had lent money to Bear Stearns did not lose their money; J.P. Morgan was able to sell off Bear Stearns' assets slowly over time, not in a fire sale, and ended up making money on the deal.

In order to make the deal work, however, the Fed had to effectively make a loan to Bear Stearns, on the collateral of Bear Stearns' assets, before J.P. Morgan ought it. It could not do that without invoking section 13(3). So it did invoke 13(3).

iii) Creation of new facilities to lend to investment banks

Once the Fed had invoked 13(3), it could set up programs ("facilities") to lend to investment banks. In March 2008 it set up two.

Primary Dealer Credit Facility (PDCF) Remember that the bond-dealing subsidiaries of investment banks were primary dealers. The PDCF made overnight repo loans to primary dealers, taking bonds as collateral.

Term Securities Lending Facility (TSLF) This lent Treasury bonds, not money, to primary dealers, taking as collateral MBS and ABS bonds (see Glossary). What was the point of this? Remember that the investment banks had been using MBS and ABS bonds as collateral to borrow overnight through repo loans. Lenders would no longer take such bonds as collaterals. But they would take Treasury bonds, and at small harcuts. So the investment banks could use the Treasury bonds lent to the by the Fed as collateral for repo loans.

iv) Lehman Brothers

In September 2008 there was a run on another investment bank, Lehman Brothers. Fed policymakers judged that this investment bank, unlike Bear Stearns, was fundamentally insolvent, not just illiquid. They did not make loans to Lehman Brothers, or make loans to support another firm's takeover of Lehman Brothers. Lehman Brothers went bankrupt.

The downfall of Lehman Brothers set off a general run on investment banks, despite the earlier creation of the PDCF and TSLF. To this day, economists continue to debate whether the Fed made a mistake about Lehman, that is whether Lehman was really insolvent or just illiquid.

v) Investment banks become banks

In response to the Lehman debacle, almost all other investment banks immediately became banks to get better access to Fed help, and to convince lenders that they wold have access to Fed help. Nowadays there are no big stand-alone investment banks; all are attached to banks. Sometimes these firms are called investment banks, but they are not investment banks in the pre-2008 sense.

d) AIG

The American International Group (AIG) was an insurance company that had got into the businenss of selling CDS (credit default swaps) on MBS (mortgage-backed securities) made out of subprime mortgages. In 2007-08 AIG had to pay out a lot of money on these. AIG started selling off its assets, which included a lot of fairly illiquid things (like companies owned by AIG). it began to look like AIG would not be able to pay off on the CDS it had entered into.

Contagion! Lots of FIs all ovr the world held the bonds AIG had sold CDS on. If AIG failed, it could hurt those other FIs. It was alreadystarting to cause runs on them.

Federal Reserve officials judged that AIG was illiquid but solvent - barely solvent, but solvent. So it made loans to AIG on the collateral of illiquid assets. (It could do this becasues it had already invoked 13(3)).

e) Money-market mutual funds

Recall that MMMFs had bought asset-backed commercial paper (ABCP) backed by subprime MBS. One of these MMMFs was called the "Reserve Primary Fund." It had been buying ABCP issed by Lehman Brothers. Uh oh! Contagion again! In September 2008 the Reserve Primary Fund announced it would "break the buck," that is it invoked its right to pay back less than 100 cents on the dollar on withdrawals - it "broke the buck." General panic, with mass withdrawals from MMMFs, ensued.

Fed policymakers wanted to do something. Their first idea was to lend to MMMFs in the way the Fed lends to banks, by making loans to an MMMF collateralized by the MMMF's assets, with haircuts. But managers of MMMFs told Fed officials that would not work. "Shareholders" (in effect depositors) in MMMFs would know that loans from the Fed were senior debt (had to be paid off before any other debts), hence had to be paid off before a MMMFs shareholders. Money borrowed from the Fed could be used to pay off withdrawals from a MMMF, but that would only increase the risk to shareholders who did not immediately withdraw. (Depositors in a bank would not worry about this sort of thing, becaise of deposit insurance.)

Simply *buying* the dodgy commercial paper from MMMFs at non-fire sale prices would do the trick. But that would violate the Fed's principle that it did not buy default-risky assets.

What to do? Fed officials designed a program called the **Asset-Backed Commercial Paper Money Market Mutual Fund Liquidity Facility (AMLF).** In the AMLF, the Fed lent to *banks* but only on one specific type of collateral: asset-backed commercial paper that the bank had *purchased from a MMMF* at a non-fire sale price. Thus, if the Fed could get banks to play along, the MMMFs would get what they needed: someone to take the asset-backed commercial paper off their hands at non-fire sale prices. To get banks to play along, the interest rate on the loans to banks had to be low, and the loans to banks had to be (of course) non-recourse.

f) Lending to issuers of commercial paper

The AMLF seemed to stop the runs on MMMFs. Burt the Fed still had a problem. "Deposits" in MMMFs were still less than they had been before the panic. And MMMFs were still wary of buying commercial paper: they preferred to hold perfectly safe Treasury bills.

That was a problem because many firms borrow by issuing commercial paper. Ordinarily a firm that usually issues commercial paper, but for some reason cannot do so in this or that month, borrows from a bank until it can issue commercial paper again. But banks were not lending in the fall of 2008 (credit crunch!). Thus, the lack of demand for commercial paper, due to the problems of MMMFs, meant that many firms were simply be unable to borrow at all. That would greatly worsen the recession that was already underway.

What to do? Fed officials designed a program called the **Commercial Paper Funding Facility (CPFF).** In the CPFF, the Fed *lent directly to non-FI businesses* by buying commercial paper they issued.

g) Swap lines

Foreign banks often take loans and deposits denominated in dollars. That is, they promise to pay the lender back in dollars, not the local currency. In the financial crisis, many foreign banks were subject to runs on their dollar-denominated borrowing. They needed dollars to pay off the withdrawn loans and deposits. How could they get these dollars? Their local central banks could make LOLR loans of the local currency, but how couuld a foreign central bank make a LOLR loan of dollars? This was a problem not just for foreign banks and foreign central banks, but also for the Fed, because financial crises in foreign countries would make the American recession worse.

The Fed's solution to this problem, created back in the 1950s, is "swap lines." A swap line is a kind of loan of dollars by the Fed to a foreign central bank, collateralized by foreign money that the foreign central bank gives to the Fed. The Fed lent a lot through swap lines starting in 2008.

The loans are called swap lines because they are structured as a "swap" of dollars for the foreign money. The Fed gives the foreign central bank dollars while the foreign central bank gives foreign money to the Fed. The foreign central bank promises to buy the foreign currency back from the Fed at a future point in time *at the same exchange rate* that was used when the Fed bought the foreign money.

The Fed *could* lose money on this deal. That would happen if the foreign central bank failed to buy the foreign currency back as promised and the foreign currency depreciated in the meantime. So the Fed only offers swap lines to a few (about nine) especially trustworthy foreign central banks including the Bank of England, the European Central Bank, and the Bank of Canada.

h) The Term Asset-backed Securities Loan Facility (TALF)

In section X) I said that a bond market, like the supply of loans to a FI, can be subject to multiple equilbria with self-fulfulling expectations. There is a good equilibrium in which people think that a type of bond will be liquid so lots of people buy the bond now; if lots of people continue to be willing to buy the bond in the future, the bond will turn out to be liquid as people expected. There is a bad equilibrium in which the bond is illiquid because people think the bond will be illiquid: few people if any buy the bond; and this continues to be true in the future so the bond turns out to be illiquid as people expected. In the good equilibrium the kind of firms that issue that kind of bond can borrow cheaply, at low yields (small liquidity premiums). In the bad equilibrium the kind of firms that issue that kind of bond borrow at high cost, at high yields (large liquidity premiums).

Central banks sometimes engage in operations meant to push a bond market toward the good equilibrium. The TALF was this kind of operation.

After Lehman Brothers went down in September 2008, people became unwilling to buy just about any kind of ABS: not just the subprime-mortage MBS that had caused all the trouble but also ABS made out

of credit card debt, student loans, car loans, loans to small businesses and commercial real estate loans - even the most highly-rated of these bonds, those rated AAA. Yields on these bonds rose so much that it was hardly worth borrowing by issuing them.

This had bad macro macroeconomic effects. The firms that issed non-MBS were mostly "finance companies." These are nonbank FIs. They make this type of loan and get money to make more loans by selling the loans they have made to companies that make ABS out of the loans and sell the bonds. Thus, the problems in the ABS market reduced the supply of loans to people borrowing through credit cards, students, car buyers and so on. It was another kind of credit crunch, not a reduction in the supply of loans from banks but rather a reduction in the supply of loans from finance companies.

To push the market out of this bad equilbrium, Fed officials designed the Term Aset-Backed Securities Loan Facility (TALF). The program started up in March 2009. In this program Federal Reserve banks would make loans on the collateral of AAA-rated non-MBS ABS bonds, with haircuts as usual. A borrower could pay the loan back any time, but could also keep the loan outstanding for a long time, as long as the typical maturity of a newly-issued ABS bond (three to five years). The loans were nonrecourse (I assume you remember what that means). TALF loans were available to just about anyone - any U.S. company.

Because TALF loans were nonrecourse, they made buying an ABS bond a better bet. Suppose it is March 2009, you buy an newly-issued ABS bond and take a TALF loan to finance part of the purchase; the other part of the purchase price, the part not covered by the TALF loan (remember the haircut) comes from your own money. There are two things that can happen: the program can succeed, making ABS bonds liquid again, or the program can fail, so that ABS bonds remain illiquid. If the program succeeds, the price of the ABS bond will rise (because the liquidity premium in its yield falls); you can sell the ABS bond, pay off the TALF loan and keep the profit. If the program fails, the price of the bond will stay low or fall; then you can walk away from the deal and keep the Fed's money. You lose the part of the bond purchase you had to finance yourself (to cover the haircut), but your loss will be limited to this.

For the same reason that the nonrecourse nature of the loan limits your possible loss, it subjects the Fed to the risk of loss: if the program fails and the price of the bond falls a lot, the value of the collateral may not be enough to cover the amount of the original loan even with a haircut. Thus, the Fed got a backstop from the Treasury for the TALF program: if there was a loss on a TALF loan because the borrower walked away from the loan and the price of the collateral had fallen a lot, the Treasury would bear the loss up to a specified amount.

As it happened, the program worked; there were practically no losses on TALF loans; there was no loss to the Tresury and the Fed made money on the program. Much more importantly, the ABS market revived, yields on ABS bonds fell; finance companies could lend again.

6) Extension of deposit insurance

This is something the FDIC did. At the time of the onset of the 2008 crisis, the FDIC insured bank deposits only up to $100,000 per deposit. That mean if you had a deposit in a bank of, say, $150,000 and the bank was closed due to insolvency, the FDIC would pay you just $100,000 and you would lose the $50,000 (unless the bank's assets turned out to be enough to cover that too). As the crisis hit and people began to fear for banks' solvency, people and businesses holding large deposits began to withdraw them from banks, at least to get them down to the $100,000 limit.

To prevent this, in October 2008 the limit was temporarily raised from $100,000 to $250,000. (The $250,000 figure was made permanent in 2010.) The FDIC also gave a temporary guarantee of all money

in the kind of checking accounts businesses use for payroll and most other transactions; these "transaction" accounts accounts do not pay interest (the program was called the Transaction Account Guarantee Program).

7) Reviving FIs: stress tests and recapitalization

In section XI) I said that after a financial crisis it can be necessary for authorities to examine the solvency of FIs, certify their solvency and recapitalize them. If private investors are not willing to take the risk of buying shares in an FI that needs more capital, a government authority might to buy shares itself, to be owned by the government agency until private investors are willing to buy them from the government agency.

FIs were examined and recapitalized in the 2008 crisis, but sort of in reverse order. The government agency that bought shares in FIs was not the Fed but the Treasury. Recall that the Fed avoids acquiring assets subject to great risk. Shares in a FI are subject to risk - that is the point of capital!

The Treasury got the money to buy shares in banks by repurposing funds that Congress had originally appropriated for another purpose. In October 2008 Congress passed a law providing money to the U.S. Treasury for an emergency program called the Trouble Asset Relief Program (TARP). The original plan was that the money would be used to buy "troubled assets," also called "toxic assets," mainly subprime MBS, from banks. As the crisis worsened, Fed and Treasury officials feared that the original plan would not help banks fast enough. They needed to convince people *immediately* that banks were solvent. So the TARP money was instead used to buy newly created shares of their stock.

The shares created to sell to the Treasury were special "preferred" stock. When a share of stock is called "preferred," that means it is privileged: any dividends paid by a company must first be paid the owners of preferred stock before anything can be paid to owners of ordinary shares ("common" stock). The idea was that the Treasury would sell its shares to private investors as soon as possible, after the crisis passed. In fact this is what happened and the Treasury made money on the program.

What about examination and solvency certification? This came later. In February 2009 the agencies that regulate banks, led by the Fed, undertook special examinations of the 19 largest banks in the country, performing stress tests on them (the first time regulators had done stress tests on U.S. banks) and reported the results to the public in detail. Some banks were required to get more capital. All of these banks succeeded in raising capital from private investors.

D) The Covid crisis

1) Introduction

In some ways, the Fed's response to the covid crisis was like its response to 2008. In fact, the Fed restarted some of the programs from the earlier crisis. But the covid crisis itself was fundamentally different from 2008, so some of the problems the Fed was trying to solve were different. Thus its response was also different in important ways. Perhaps the most important innovation: lots more lending to non-FI businesses, and even to local governments.

2) New constraint on Fed lending to nonbanks

The constraints on Fed operations were also a bit different from what they had been in 2008. In section XI) I mentioned the Dodd-Frank Act of 2010. One of the things in the Dodd-Frank Act was an

amendment to section 13(3) of the Federal Reserve Act, the section that allows the Fed to lend to nonbanks under extreme circumstances. The amendment placed a new restriction on the way the Fed can do this.

Before Dodd-Frank, after the Fed had declared an emergency, it could arrange a loan with an individual nonbank tailored to that firm's circumstances. In 2008 the Fed made such individually-tailored loans to Bear Stearns and AIG. Under the amendment made by Dodd-Frank the Fed can lend to nonbanks only through "broad-based facilities" available to an entire class of nonbanks. "Broad-based" means that there must be at least five firms that are eligible to borrow under the rules of the facility. The things that the Fed did in 2008 that were called "Facility" would all still be OK under the new rules. But the loans that the Fed made to Bear Stearns and AIG would not.

3) The crisis

a) Similar to 2008

When Covid hit in March 2020 there was a general withdrawal of lending to FIs. Due to this, and to a desire of FIs to stay liquid, there was a credit crunch - a decrease in the supply of loans from FIs. There was also a decrease in demand for commercial paper, and mass withdrawals from MMMFs that buy commercial paper. All of these things were similar to 2008.

b) Different from 2008

i) Withdrawal of "term lending"

In 2008 the withdrawal of lending to FIs was largely a withdrawal of overnight lending, due to fear that a borrower would go bankrupt (standard liquidity-crisis stuff). In 2020 it was rather a withdrawal of "term lending" (also called "term funding'). These are loans that are short-term, but longer than overnight, like a week or a month. The withdrawal of term funding was due to uncertainty about the near future - an unwillingness to commit funds longer than overnight.

ii) Increase in need to borrow by non-FI businesses and local governments

Covid shutdowns rapidly decreased firms' sales; firms needed to borrow to cover rent, payrolls and other continuing expenses. State and local governments needed to borrow a lot more for Covid-related programs.

iii) Treasury and agency MBS markets become less liquid

In 2020 there "problems with liquidity" in the markets for agency MBS and even Treasury bonds. Ordinarily Treasury bonds are as liquid as anything can be: there are always dealers ready to buy these types of bonds at low bid-ask spreads (see section VI). In March 2020 there were so many people trying to sell Treasury and agency MBS bonds that dealers couldn't handle them all. Dealers' inventories were filling up. There is a limit on how much a bond dealer can hold in inventory. Remember dealers have to borrow to finance their inventories. They can borrow only so much, because people are willing to lend only so much to any one dealer. So some dealers cut back on buying; some stopped buying altogether; bid-ask spreads went up a lot.

Nothing like this had happened in 2008. Then, demand for Treasury bonds actually increased. Demand for "agency MBS," that is MBS issued by the GSEs (Fannie Mae and Freddie Mac), made out of conforming (prime) mortgages, held up pretty well despite the crisis in the subprime MBS market.

4) What the Fed did

a) Restart programs from 2008

These programs were restarted:

Primary Dealer Credit Facility (PDCF)
Commercial Paper Funding Facility (CPFF)
Term Asset-backed Securities Loan Facility (TALF)
Swap lines (providing dollars to foreign central banks)

The Asset-Backed Commercial Paper Money Market Mutual Fund Liquidity Facility (AMLF) was restarted with a change, and a slightly different name: the new program was called the Money Market Mututal Fund Liquidity Facility (MMLF). The old 2008 program had lent money to banks to buy asset-backed commercial paper from MMMFs. The new 2020 program lent money to banks to buy any kind of commercial paper (and even Treasury bills) from MMMFs.

b) To help the Treasury bond market

i) Open market operations in Treasuries

Recall that the problem in the Treasury and agency MBS markets was lots of people attempting to sell, and dealers who couldn't buy because they couldn't get repo loans to finance larger inventories.

To help with this, the Fed did something it often does in open-market operations. It bought lots of Treasury bonds and agency MBS bonds. It made overnight and term repo loans to primary dealers, collateralized by Treasuries and agency MBS.

ii) Foreign and International Monetary Authorities Repo Facility (FIMA)

Recall that foreign central banks sometimes need dollars to lend in dollars to their local banks that have borrowed or taken deposits in dollars. Foreign central banks with which the Fed has established swap lines can borrow dollars through swap lines. But this is only a few central banks (about nine). What about other central banks? Ordinarily, they hold large amounts of U.S. Treasury bonds. When they need dollars, they sell some of those Treasury bonds in the U.S.

Many foreign central banks needed dollars in spring 2020. They were selling lots of bonds in the U.S. The Fed didn't want them selling even more masses of Treasury bonds in U.S. markets -that would only worsen the problems in those markets. But the Fed didn't want to grant swap lines to every Tom, Dick and Harry. So the Fed established the Foreign and International Monetary Authorities Repo Facility (FIMA). This was a facility that would make repo loans of dollars to foreign central banks, taking Treasury bonds as the collateral.

c) Lending to non-FI firms and governments

In 2008 the Fed had established a facility that lent to entities that were not banks, and not even FIs: the Commercial Paper Funding Facility (CPFF). Through the CPFF, the Fed in effect lent to any companies that issued commercial paper. The TALF would also lend to non-FI businesses that bought ABS bonds.

In 2020 the Fed established a *lot more* facilities to lend to a wide variety of non-FI businesses and even local governments. All had backstops of funding from the Treasury to limit risk to the Fed.

Primary Market Coporate Credit Facility (PMCCF) bought newly issued corporate bonds (and "syndicated loans," a sort of bond-loan hybrid)

Secondary Market Corporate Credit Facility (SMCCF) bought the same things but not newly-issued; bought old ones from investors. It could also purchase ETFs that held crporate bonds. An ETF is a mutual fund whose shares are traded on a stock exchange. When you buy a share in the ETF, you are essentially buying a fraction of the portfolio of assets held by the ETF.

Main Street Lending Program (MSLP) bought from banks loans the banks had made to small businesses

Municipal Liquidity Facility (MLF) lent to states, counties and cities.

E) The March 2023 crisis

1) Introduction

The financial crisis of 2008 and the incipient crisis of spring 2020 were both mainly about liquidity. The incipient crisis of spring 2023 was different: it was about interest-rate risk.

So far Federal Reserve policymakers and staff have said very little about how their response to the spring 2023 crisis fits in with generally-accepted principles of central banking such as Bagehot's rules. So I will not have anything to say about that in these notes.

2) The crisis

Coming out of the 2023 Covid crisis, unemployment quickly fell to low rates and inflation picked up rapidly. In response the Fed raised the general level of interest rates. (I will explain how this works, and why the Fed would want to raise interest rates in such a situation, in later sections of the course.) Figure 2 at the end of this section shows the monthly-average yield on ten-year Treasury coupon bonds starting in June 2009, around the end of the recession that followed the 2008 financial crisis, and ending in July 2023. Notice the big increase in the yield beginning in late 2020. The general increase in bond yields meant a decrease in the market values of bonds. Values of longer-term bonds fell the most, of course.

The **Silicon Valley Bank (SVB)** was a "mid-sized" bank in California. It was the sixteenth-largest bank in the U.S. (measured by assets). It specialized in dealing with tech firms. In 2020-21 many of its tech firm customers got a lot of money from venture capital investors. But the tech firms really did not have anything productive to do with the money, so they just deposited it in the SVB. Many of the deposits were large, in excess of the $250,000 limit on FDIC deposit insurance.

Meanwhile, there was little demand for loans from SVB: for the same reason its tech firm customers did not have anything productive to do with the money they got from venture capitalists, they had no use for loans from SVB. So SVB put the money into Treasury bonds. Yields were higher on long-term Treasury bonds, of course, due to term premiums. So SVB bought long-term bonds, especially.

When yields rose over 2022, the market value of SVB's long-term Treasury bonds fell a lot. In early 2023 people began to fear that SVB might be getting close to insolvency. At the beginning of March 2023 there were mass withdrawals of uninsured deposits in SVB - a run. Regulators closed the bank down.

The run on SVB and its closing made people ask whether other banks might be in a similar situation. **Signature Bank** was another mid-sized bank, the twenty-ninth largest bank in the country, with a lot of deposits in excess of the $250,000 limit. It was headquartered in New York city but had a lot of offices in California and had made a specialty in dealing with crypto-related firms. The day SVB was closed by regulators, there was a run on Signature Bank, which was then closed by regulators.

First Republic Bank was yet another mid-sized bank, fourteenth-largest in the U.S., headquartered in San Francisco. It had a lot of deposits in excess of $250,000, partly because it specialized in dealing with rich people. In response to the SVB and Signature Bank closings, people began to withdraw deposits from First Republic. To boost confidence in First Republic and stop the panic from spreading, several very large banks deposited a lot of money in First Republic a few days after Signature Bank was closed. The withdrawals continued, however, and at the beginning of May regulators closed the bank.

3) What the Fed did: the Bank Term Funding Program (BTFP)

In March 2023, just after the closing of SVB and Signature, the Fed announced the creation of a new facility, the Bank Term Funding Program (BTFP). The BTFP would lend to banks on the collateral of Treasury and MBS bonds issued by Fannie and Freddie. The amount of the loans would be equal to the par value of the bonds, not their current market value which was well below par value because of the recent big increase in bond yields. Also, no haircuts would be applied. This was a big change from previous Fed practice. In previous Fed lending, bonds used as collateral for a loan were valued at current market prices, with haircuts.

Lending this way could expose the Fed to a lot of risk if a borrower defaulted, because the market value of the collateral would be less than the amount of the loan. Thus, the Fed got a Treasury backstop for the facility.

4) What the FDIC did: pay off uninsured deposits

In March 2023, just after the closing of SVB and Signature Bank, the FDIC announced that all deposits in the two banks, not just those up to the $250,000 limit, would be paid off by the FDIC's deposit insurance fund. The extra cost to the FDIC's insurance fund would be covered by a special assessment on banks. The 1991 law governing the FDIC allows the FDIC to do this sort of thing by declaring a "systemic risk exception."

Table 1

FRED — Real Residential Property Prices for United States
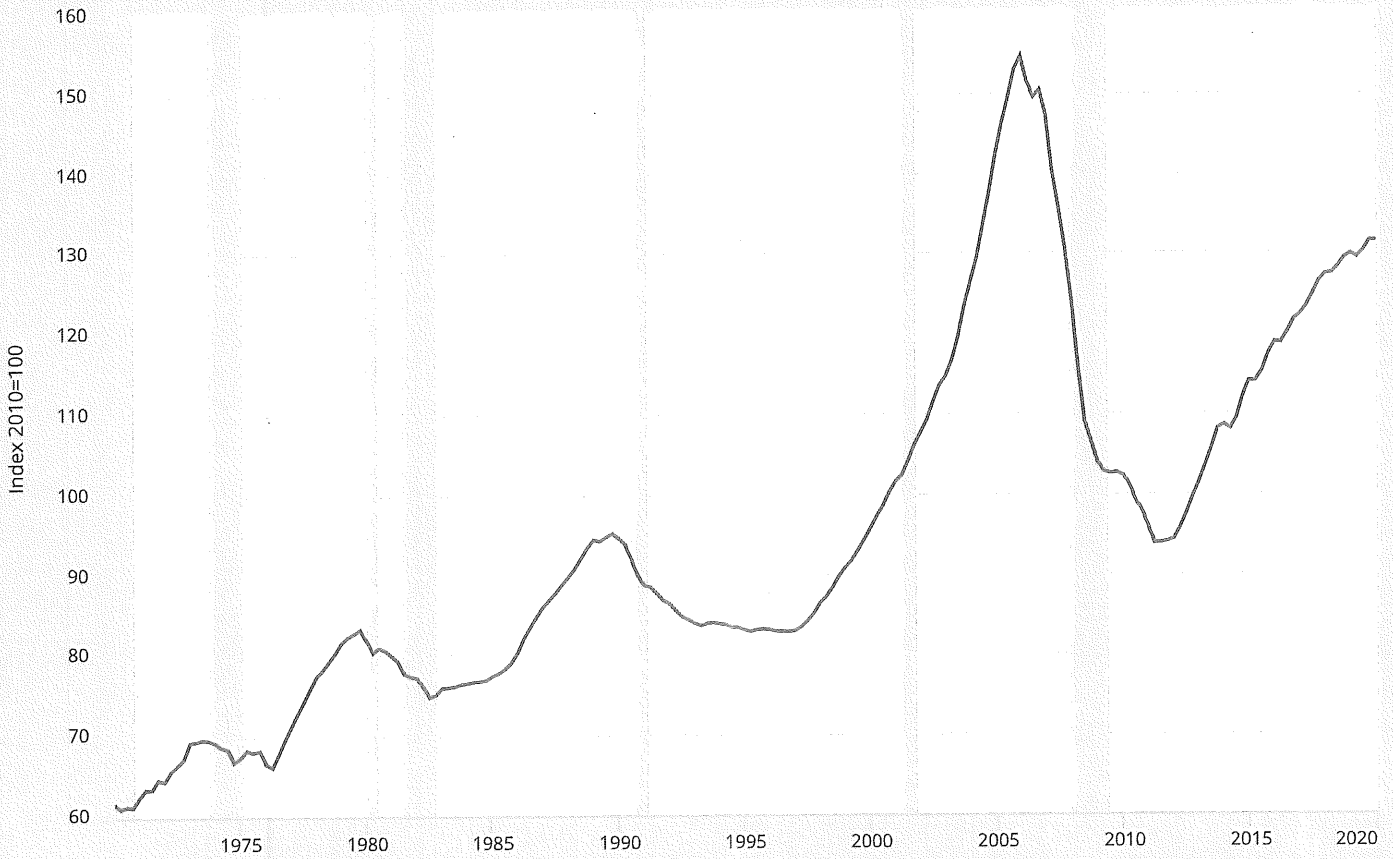
Source: Bank for International Settlements   fred.stlouisfed.org

Table 1

Table 2

FRED — Market Yield on U.S. Treasury Securities at 10-Year Constant Maturity, Quoted on an Investment Basis

Percent

4.5

4.0

3.5

3.0

2.5

2.0

1.5

1.0

0.5

2012   2014   2016   2018   2020   2022
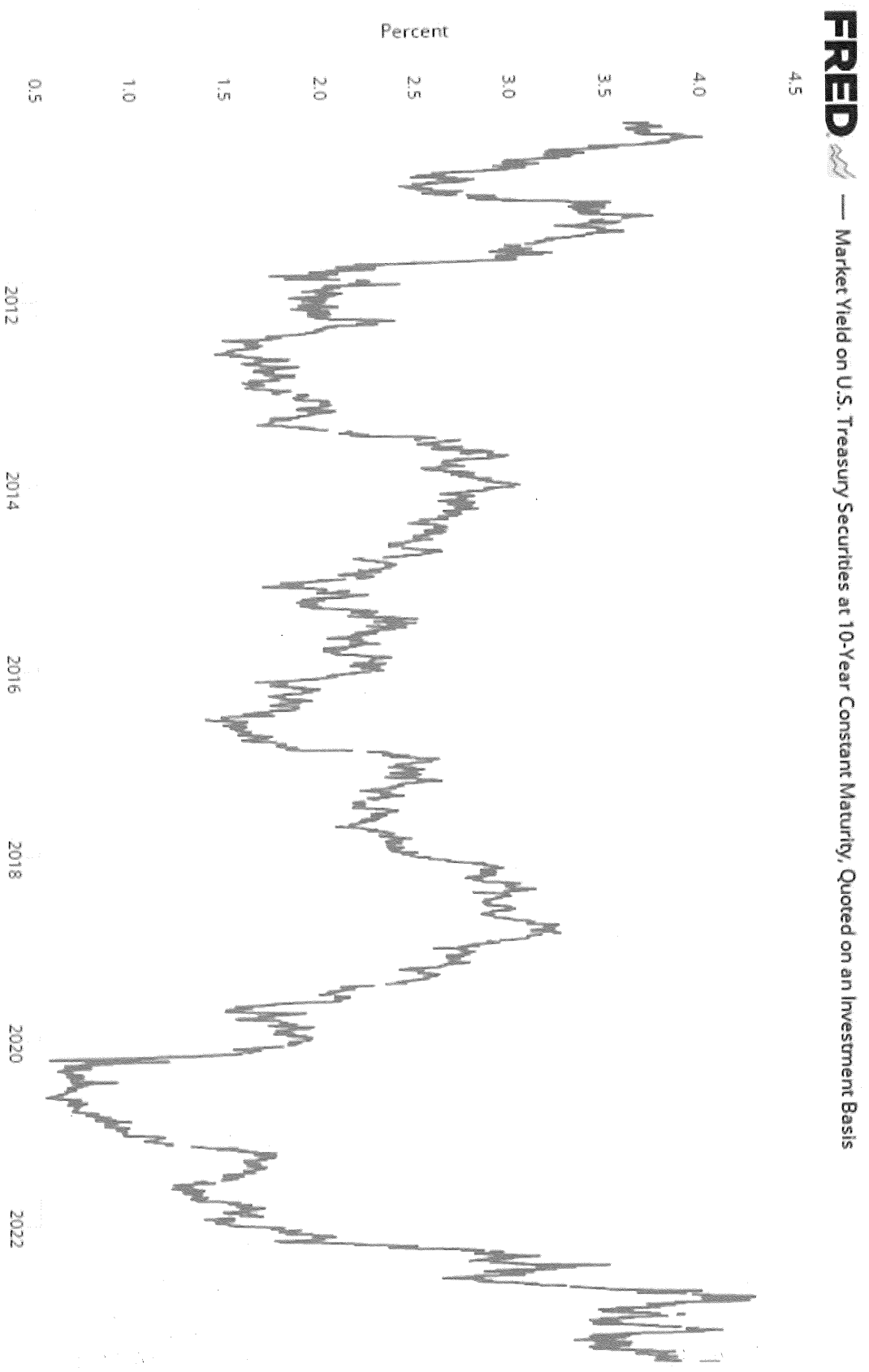
Source: Board of Governors of the Federal Reserve System (US)

myf.red/g/16XN5

©**Christopher Hanes**   **Economics 450   Class notes**

XIII) How a central bank controls the overnight interest rate
    A) Introduction and preview
    B) Reserve supply
        1) Preview
        2) Reserve accounts and reserve supply
        3) Payments between banks do not affect reserve supply
        4) U.S. Treasury payments affect reserve supply
        5) Changes in the public's demand for cash affect reserve supply
        6) Federal Reserve lending can affect reserve supply
        7) Federal Reserve open-market operations affect reserve supply
    C) Reserve demand
        1) Introduction
        2) Analogy: your checking account
        3) Opportunity cost of leaving funds in reserve account
            a) Market overnight rate
            b) Types of overnight loans
                i) Fed funds
                ii) Overnight repurchase agreements (repos)
                iii) Overnight eurodollars
        4) Interest on reserve accounts
        5) Penalty for overdraft/emergency loans
            a) What happens if a bank overdraws its reserve account
            b) Primary credit (discount loans)
        6) Reserve requirements
        7) Demand for reserve balances
            a) What it looks like
            b) Why perfectly elastic demand at r_P , r_D ?
                i) r_D
                ii) r_P
            c) Changes in r_P , r_D
            d) Changes in reserve requirements
    D) Systems of control
        1) Introduction
        2) Floor system
        3) Corridor (or channel) system
        4) Fed's old pre-2007 system
        5) Complication in the U.S.: subfloor
            a) In the U.S., r actually *can* fall below r_D
            b) Some holders of reserve accounts don't get paid interest on reserves
            c) Solution: subfloor (ON RRP)
    E) Mathematical model of reserve demand
        1) Introduction
        2) Review: things from this class
            a) Liquidity preference
            b) Using calculus to find the optimal value of a choice variable.
        3) Review: ideas from stats
            a) The distribution of a random variable and its histogram
            b) The c.d.f. of a random variable
            c) Conditional expectation (or expected value)

d) The uniform distribution
  i) Definition
  ii) Formulas
4) Liquidity preference in the context of reserve demand
5) The model
  a) Story and notation
  b) What a bank wants to maximize
  c) Uniform distribution for P
  d) Reserve demand

XIII) How a central bank controls the overnight interest rate

A) Introduction and preview

In this section I explain how a central bank controls the overnight rate. When we say "the overnight rate" we mean the average rate for overnight loans negotiated between borrowers and lenders. A central bank does not control this directly - it does not issue a command that "all overnight loans today must be at a rate of 1.25 percent." Rather, a central bank controls the rate indirectly by controlling supply and demand in a market.

As an analogy, consider a perfectly competitive market for fish in a coastal fishing village. For simplicity, we will say there is just one kind of fish. At the end of every day fishermen bring their catch to the market while shoppers for restaurants and families come to buy. There is a demand curve for fish. The daily supply of fish is fixed (equal to the amount the fishermen caught that day). There is a market-clearing equilibrium price for fish.
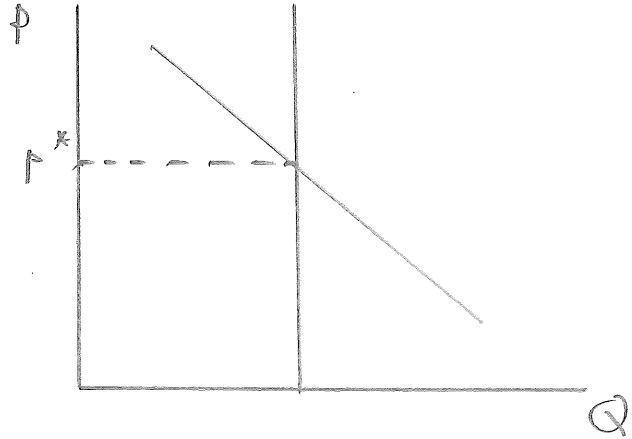
Suppose you want to control the price of fish: you want to be able to pick a number and make the market price equal that number. You could do this if you could manipulate the supply of fish: then, taking the position of the demand curve as given, you could make the supply equal the value that makes the market price hit your target number. Alternatively, you could hit your target number by controlling the position of the demand curve, taking supply as given.

The market overnight rate is, like the price of fish, a market-clearing equilibrium price resulting from the interaction of a supply and a demand. You might think that I am talking about the supply and demand for overnight loans. In a sense that is true, but you will see that it is easier to understand what is going on by thinking about supply and demand for *reserves*. Recall that "reserves" means money held by banks. In an economy with a central bank, "reserves" means money banks hold in their accounts at the central bank. A central bank controls both the supply of reserves and the demand for reserves. Thus, it can control the market overnight rate. Supply and demand in the market for reserves looks like this: or this:
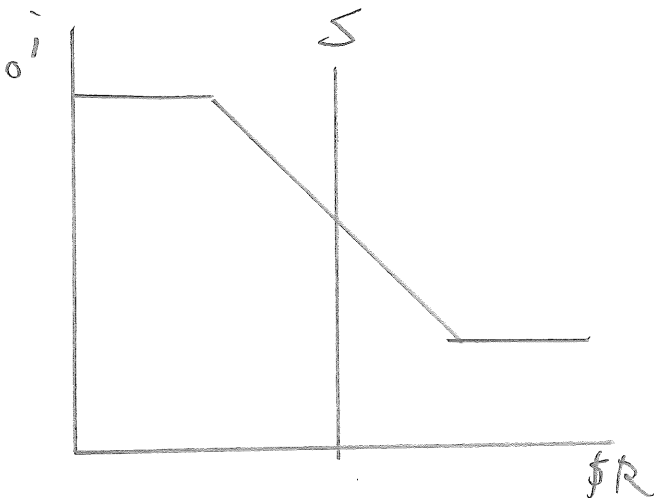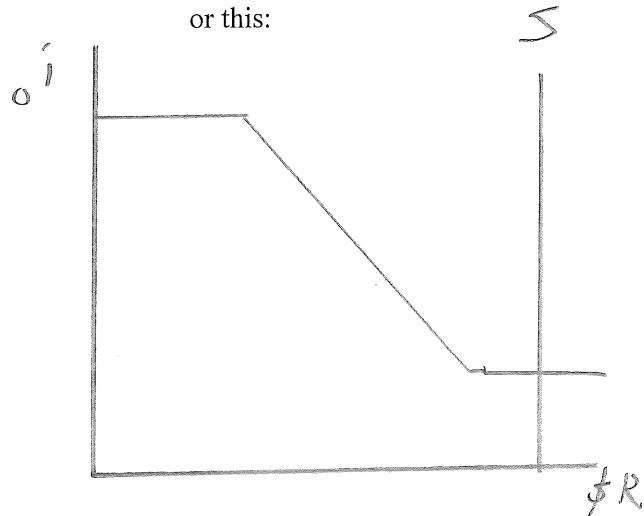
Note that the supply of reserves is fixed, like the supply of fish in our village fish market. The demand curve looks odd. I'll start by explaining reserve supply and then move on to demand.

B) Reserve supply

1) Preview

Reserve supply is total dollars in all banks' Federal Reserve accounts. Payments between banks do not affect reserve supply. Reserve supply *is* affected by:
- payments made or received by the U.S. Treasury
- changes in the public's demand for cash
- Some types of loans made by the Federal Reserve
- Federal Reserve open-market operations

2) Reserve accounts and reserve supply

Recall that banks use money to make payments to other banks. Say you pay your rent with a check for $500. Your landlord deposits the check in his bank. Recall that means your landlord asks his bank to collect the payment and add it to his deposit at his bank. Your landlord's bank presents the check to *your* bank and says "pay me money!"

Banks also use money to deal with payments to and from the government. Say you are a government employee. You deposit a $1,000 paycheck in your bank. The government must pay that amount of money to your bank. Say your parents pay $2,000 in income tax with a check to the Internal Revenue Service. The U.S. Treasury gives the check to your parents' bank and says "pay me money!"

Banks use their reserve accounts in the Federal Reserve system to make and receive all these types of payments. The U.S. Treasury also has an account in the Federal Reserve system. The Treasury's account at the Fed is not called a reserve account, but otherwise it is the same.

You can think of banks' reserve accounts and the U.S. Treasury's Fed account as numbers on an Excel spreadsheet on a Federal Reserve computer. To keep things simple, I will pretending there are just three banks and I'll keep the numbers small. Like this:

Federal Reserve's Excel spreadsheet

Bank	$ in account
Citibank	100,000
HSBC	85,000
Chemung Canal	12,000
Total Reserves	197,000
U.S. Treasury	45,000

"Reserve supply" is the total of dollars in all banks' reserve accounts. In this example reserve supply is $197,000. The Treasury's balance is *not* part of reserve supply.

3) Payments between banks do not affect reserve supply

Consider your $500 rent payment to your landlord. Say your bank is HSBC, and your landlord's bank is Citibank. HSBC pays Citibank $500 by sending the Fed a message: "Take $500 out of the HSBC reserve account, transfer it to Citibank's reserve account." After the Fed has changed its Excel spreadsheet accordingly, it looks like:

Bank	$ in account
Citibank	100,500
HSBC	84,500
Chemung Canal	12,000
Total Reserves	197,000
U.S. Treasury	45,000

Reserve supply is unchanged. Payments between banks move numbers around from one row to another, but they do not affect reserve supply.

4) U.S. Treasury payments affect reserve supply

Consider your $1,000 paycheck from your Federal government job. The Treasury pays your bank, HSBC, by sending the Fed a message: "Take $1,000 out of the Treasury Fed account, transfer it to HSBC's reserve account." After that payment (*and* the $500 payment from HSBC to Citibank) the Fed's Excel spreadsheet would look like this:

Bank	$ in account
Citibank	100,500
HSBC	85,500
Chemung Canal	12,000
Total Reserves	198,000
U.S. Treasury	44,000

Reserve supply has increased by $1,000.

And after your parents' tax payment of $2,000, assuming their account is at the Chemung Canal Bank,

Bank	$ in account
Citibank	100,500
HSBC	85,500
Chemung Canal	10,000
Total Reserves	196,000
U.S. Treasury	46,000

Reserve supply has decreased by $2,000.

5) Changes in the public's demand for cash affect reserve supply

A lot of cash - paper money and coins - is held in American pockets, cash registers, ATMs, and so on. A lot of American cash is held *outside* the United States. Cash can pass from hand to hand without affecting the total amount in circulation. But there are also changes in the public's total demand for cash. If the public wants to hold more cash, how do they get it? By withdrawing cash from their bank accounts. Where do banks get cash to give their customers? By withdrawing from their reserve accounts! That reduces reserve supply.

Suppose customers of the Chemung Canal bank withdraw $800 in cash from their accounts, leaving the Chemung Canal bank's ATMs empty. The Chemung Canal Bank withdraws $800 from its reserve account to replenish its ATMs. Then the Fed's Excel spreadsheet looks like:

Bank	$ in account
Citibank	100,500
HSBC	85,500
Chemung Canal	9,200
Total Reserves	195,200
U.S. Treasury	46,000

Reserve supply has decreased by $800. A decrease in the public's demand for cash has the opposite effect, of course.

6) Federal Reserve lending can affect reserve supply

The Federal Reserve makes loans of many types. For historical reasons, all types of loans made by the Federal Reserve system are sometimes referred to as "discount loans." It is said that discount loans are made "through the discount window." In fact, there are distinct names for the different types of loans the Fed makes. One is "primary credit." This is a short-term loan to a bank meant to cover an overdraft or other type of accidental shortfalls in the bank's reserve account. Dollars lent by the Fed for this purposes do not count as part of reserve supply, as I will explain later.

Dollars lent by the Fed for *other* purposes *are* part of reserve supply. Thus, reserve supply increases when the Fed makes a loan other than a primary credit loan. Reserve supply decreases when the borrower pays off such a loan. As an example, suppose that the Fed makes a $1,000 loan to a company that has an account at HSBC. The Fed gives the company a check which the company deposits in its HSBC account. HSBC presents the check to the Fed and says "pay me." The Fed does that by adding $1,000 to HSBC's reserve account. Then the Fed's Excel spreadsheet looks like:

Bank	$ in account
Citibank	100,500
HSBC	86,500
Chemung Canal	9,200
Total Reserves	196,200
U.S. Treasury	46,000

Reserve supply has increased by $1,000.

7) Federal Reserve open-market operations affect reserve supply

The Federal Reserve buys and sells bonds, mainly Treasury bonds but also some other types of bonds. When the Fed buys or sells bonds, that is called "open market operations."

When the Fed buys anything, including bonds, reserve supply increases.
When the Fed sells anything, inccluding bonds, reserve supply decreases.

As an example, suppose that the Fed buys a bond from you for $800. The Fed gives you a check. You deposit the check in your account at HSBC. HSBC takes the check to the Fed and says "pay me." The Fed does that by adding $800 to HSBC's reserve account. Then the Fed's Excel spreadsheet looks like:

Bank	$ in account
Citibank	100,500
HSBC	87,300
Chemung Canal	9,200
Total Reserves	197,000
U.S. Treasury	46,000

Reserve supply has increased by $800.

Make up your own example of the Fed *selling* something. I expect you have the idea by now. I am too bored to go on.

C) Reserve demand

1) Introduction

Above I said that reserve demand looks like:
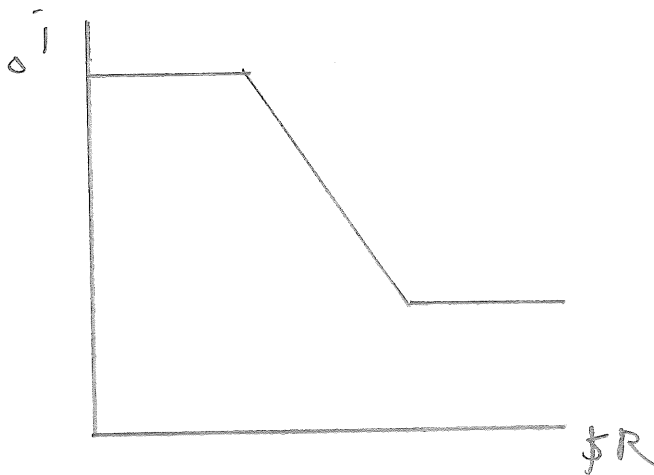
In this section I will explain why. There are two things to explain.
- Why is reserve demand flat at the top and bottom?
- Why is reserve demand downward sloping in between the top and bottom? That is to say, why is the quantity of reserves demanded by banks negatively related to the makret overnight rate?

7

2) Analogy: your checking account

The nature of reserve demand reflects the rules that the Fed applies to banks' reserve accounts. These rules are similar to those governing your checking account at a bank. Thus, to understand reserve demand it helps to think about your checking account.

Opportunity cost. Remember that an opportunity cost is what you give up when you choose to do something. There is an opportunity cost to you associated with the funds you leave in your checking account. . If you leave $500 in your checking account on average, that is $500 less you can put into bonds, or mutual funds, or whatever your next-most-liquid investment would be. Thus, the opportunity cost of leaving money in your checking account is the interest you could earn on your next-most-liquid investment.

Interest on your checking account. Some checking accounts (perhaps not yours) pay interest.

Overdrafts. If you run an overdraft in your checking account, you are subject to some kind of penalty. For many accounts it's a fee, like $30.

Uncertainty about what your balance is at any point in time. You can't predict exactly what the balance will be in your account at the end of any given day, because you don't know *exactly* when payments into the account (such as checks you deposit) will be credited, or payments out (checks you write, debit card purchases) will be debited.

Minimum balance requirement. Many checking accounts have a minimum average balance requirement. For example, if the balance in the account on average over a month falls below $500, you are subject to some kind of penalty.

3) Opportunity cost to a bank of leaving funds in its reserve account

a) Market overnight rate

The opportunity cost to a bank of leaving funds in its reserve account is the market overnight rate. Every dollar that a bank leaves in its reserve account overnight is a dollar the bank did not lend overnight. Of course, every dollar that a bank leaves in its reserve account is also a dollar it did not use to lend for a lonegr period, or to buy bonds. Why focus on overnight loans? Because an overnight loan is the asset that is the closest substitute for a dollar left in the reserve account. An overnight loan is perfectly liquid and not subject to interest-rate risk. (Treasury bills are also perfectly liquid, but they are subject to a little bit of interest-rate risk.)

In this section I will use r to denote the market overnight rate. I know this is confusing, sorry. Up to now I have used $_o i$ to denote the market overnight rate. I am switching to r here to match the notation in the coursepack reading about this topic.

b) Types of overnight loans

i) Fed funds

This is an overnight loan without collateral between entities that both have accounts in the Federal Reserve system. Usually banks, but also some government-related financial agencies you will learn about later. (The U.S. Treasury has an account at the Fed, but it does not engage in fed funds lending.) Frequently, people use the term "fed funds rate" to refer to the interest rate on overnight loans in general, because on any day all overnight interest rates are very nearly the same.

ii) Overnight repurchase agreements (repos)

An overnight loan collateralized through a repurchase agreement (q.v.). All kinds of businesses, including but not limited to FIs, borrow and lend overnight through repurchase agreements.

iii) Overnight eurodollars

An overnight deposit, not collateralized, in a bank outside the U.S., in dollars not the bank's home currency. U.S. banks are not allowed to pay interest on this type of deposit, so when U.S. banks borrow dollars overnight it is through fed funds or repo.

4) Interest on reserve accounts

The Fed pays banks interest on their reserve accounts. I denote this interest rate as r_D to match the coursepack reading.

5) Penalty for overdraft/emergency loans

a) What happens if a bank overdraws its reserve account

A bank can overdraw its reserve account just as you can overdraw your checking account. At the end of a business day, the Federal Reserve turns off "Fedwire." Fedwire is the communications network that banks and other holders of Federal Reserve accounts use to make payments between those account. Then the Fed finishes settling the payments for the day. If a bank ends up with a negative balance, the account is overdrawn.

b) Primary credit (discount loans)

If an account is overdrawn when the Fed finishes settling all the payments for the day, the Fed automatically makes an overnight loan to the bank to cover the shortfall. The type of loan from the Fed used to cover overdrafts in banks' reserve account is called "primary credit." The rate charged is called the "primary credit rate." Thus, the cost to a bank of running an overdraft in its reserve account is the primary credit rate. I will denote the primary credit rate as r_P, again to match the reading . It is always higher than r_D .

When we say "reserves" or "reserve supply," the number we are thinking of does *not* include the dollars lent by the Fed to cover overdrafts in banks' reserve accounts. The absolutely correct phrase that means "total dollars in banks' reserve accounts not including loans to cover overdrafts" is "nonborrowed reserves." When I say "reserves" below, I mean, to be quite correct, "nonborrowed reserves."

Remember that people call all kinds of loans from the Fed, informally, "discount loans." So sometimes, confusingly, primary credit loans are called "discount loans" and the primary credit rate is called the "discount rate."

6) Reserve requirements

In the past, banks' reserve accounts at the Fed were subject to minimum balance requirements called "reserve requirements." A bank had to hold a certain number of dollars in the account on average over *two week* periods called "maintenance periods." At the end of the last day of a maintenance period, after the Fed has settled all the payments for the day, if a bank had not covered its reserve requirement for the maintenance period, it could borrow at the primary credit rate to cover the shortfall. The dollar amount of a bank's reserve requirement for a maintenance period is equal to ten percent of the amount of money in checking accounts at that bank, as reported by the bank to the Fed, on average over the two weeks prior to the maintenance period.

The Fed eliminated reserve requirements in March 2020.

7) Demand for reserve balances

a) What it looks like

This is what reserve demand
looks like:
See that reserve demand is
"perfectly elastic" (flat) at
r_P and r_D, and
downward-sloping in between.
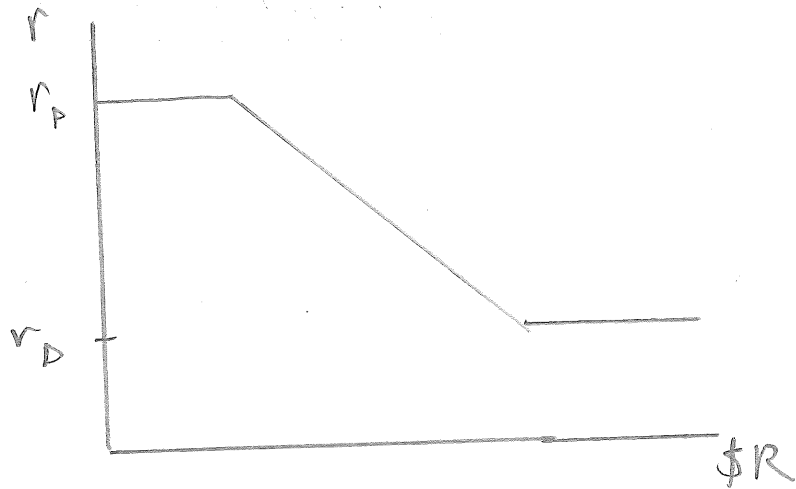
Putting reserve supply together with reserve demand, the market overnight rate depends on what reserve supply is, relative to the position of the reserve demand curve:
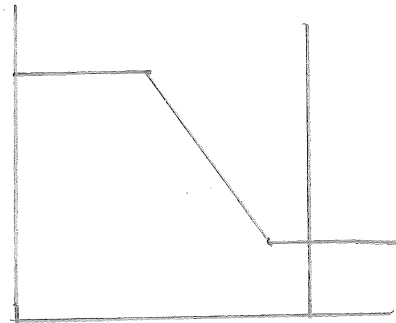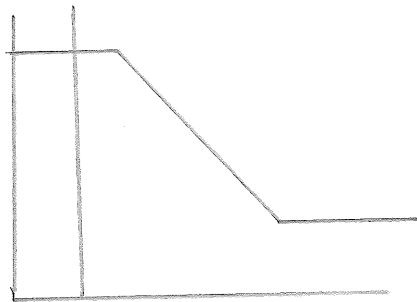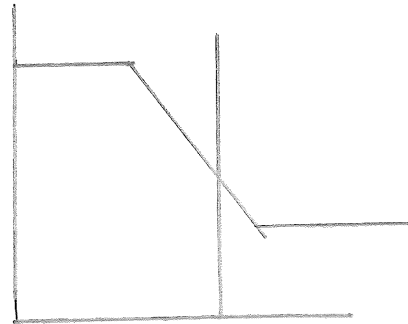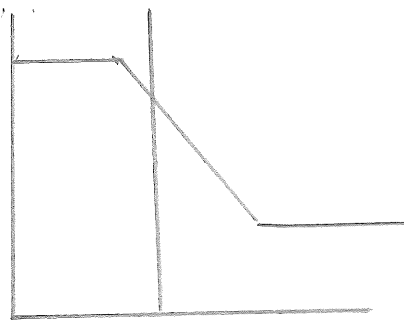
Notice that the market overnight rate can't go higher than r_P, or lower than r_D.

b) Why perfectly elastic demand at r_P, r_D ?

i) r_D

Why can't the market overnight rate fall below r_D ?

Well, just by leaving funds in its reserve account, a bank can "lend" to the Fed and get an interest rate of r_D. So no bank would lend to anyone else at a rate lower than r_D! So the market rate can't go lower than that. (There's a complication here that we'll get back to later.)

The flatness of the reserve demand curve at r_D indicates that, when the market overnight rate is equal to r_D, banks want to hold a quantity of reserve balances *at least* equal to $\hat{R}$, but don't really care how much they hold as long as it is at least $\hat{R}$.
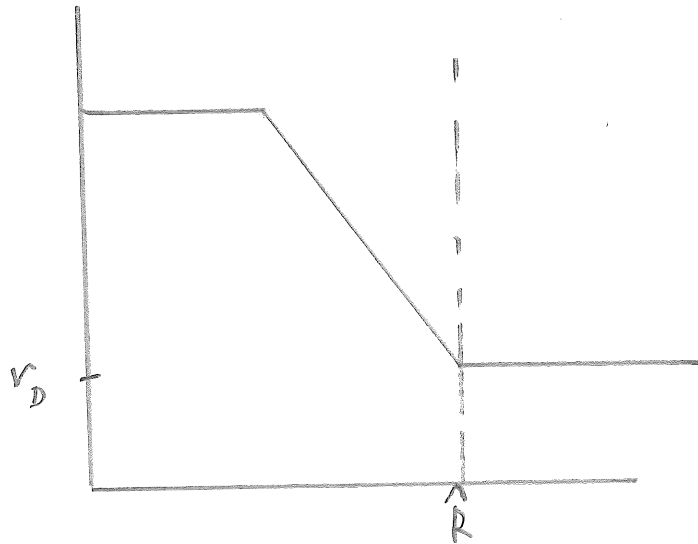
Why is that? If the market overnight rate is equal to r_D, then the opportunity cost of holding another dollar in the reserve account is exactly equal to the interest paid on the reserve account. So hold another dollar, don't hold another dollar - who cares?

ii) r_P

Why can't the market overnight rate rise above r_P?

Well, just by intentionally overdrawing its reserve account, a bank can borrow from the Fed at an interest rate equal to r_P. So no bank would pay more than that to borrow in the market.

The flatness of the reserve demand curve at r_P indicates that, when the market overnight rate is equal to r_P, banks want to hold a quantity of reserve balances *no bigger* than $\hat{R}$, but don't really care how much they hold as long as it is less than or equal to $\hat{R}$.
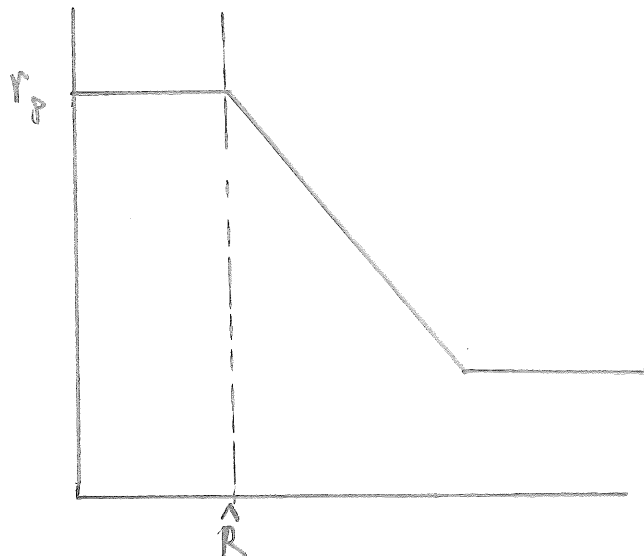
Why is that? Consider the cost and benefit of intentionally overdrawing the reserve acccount by another dollar. The cost is r_P. The benefit is that I have another dollar to lend on the market and earn the market overnight rate r. If $r = r_P$, who cares how big my overdraft is?

c) Changes in r_P, r_D

If the Fed increases both r_P and r_D by the same amount then the reserve demand curve slides up. Say the Fed increase both rates by 1/2 percent:
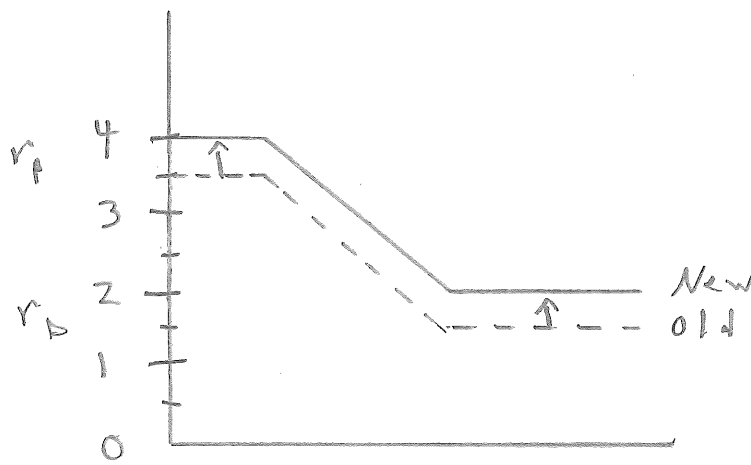
The opposite holds for decreases.

d) Changes in reserve requirements

An increase
in the dollar
amount of required
reserve balances
slides the reserve
demand curve to
the right:
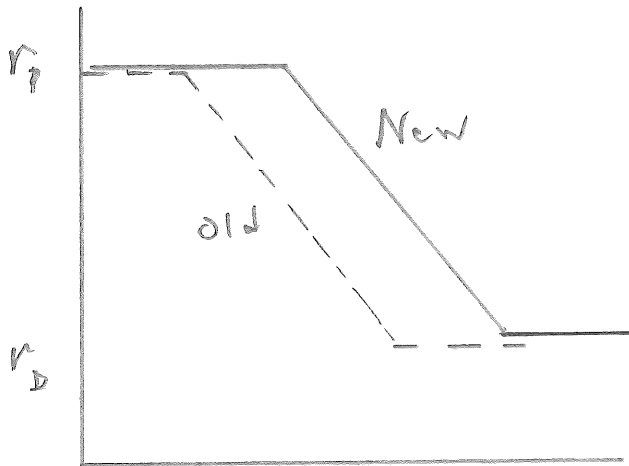
A decrease slides it to the left.

7) Reserve supply and the market overnight rate

The market overnight rate depends on the configuration of r_P, r_D and (nonborrowed) reserve supply. The Fed directly controls all of these things. So it can *indirectly* control the market overnight rate and make the overnight rate hit the target (often called the target fed funds rate) chosen by the FOMC. To match the r's, I will denote the target by r_T.

8) Systems of control

a) Introduction

Central banks have used different methods to hit their targets. Right now, the Federal reserve and most other central banks use the "floor" system. At the time the 2008 financial crisis hit, most were using the "corridor" (or "symmetric corridor") system, also called the "channel" (or "symmetric channel") system. Prior to 2007, the Fed used yet another system.

b) Floor

In the floor system, the central bank sets $r_D = r_T$ and keeps reserve supply high enough to ensure that the market rate stays on the floor, equal to the rate of interest the central bank pays on reserve balances.
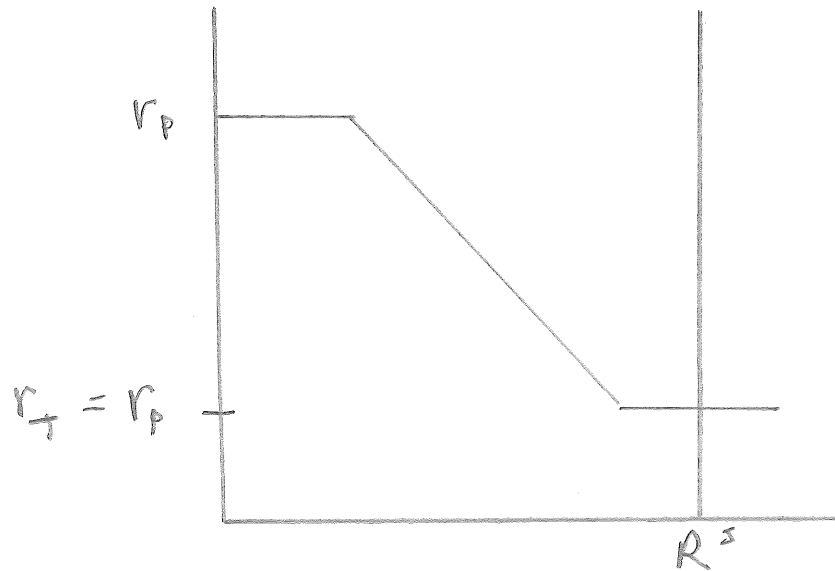
When the central bank policy committee raises (lowers) r_T , it just raises (lowers) r_D and r_P .

Increase in r_T Decrease in r_T
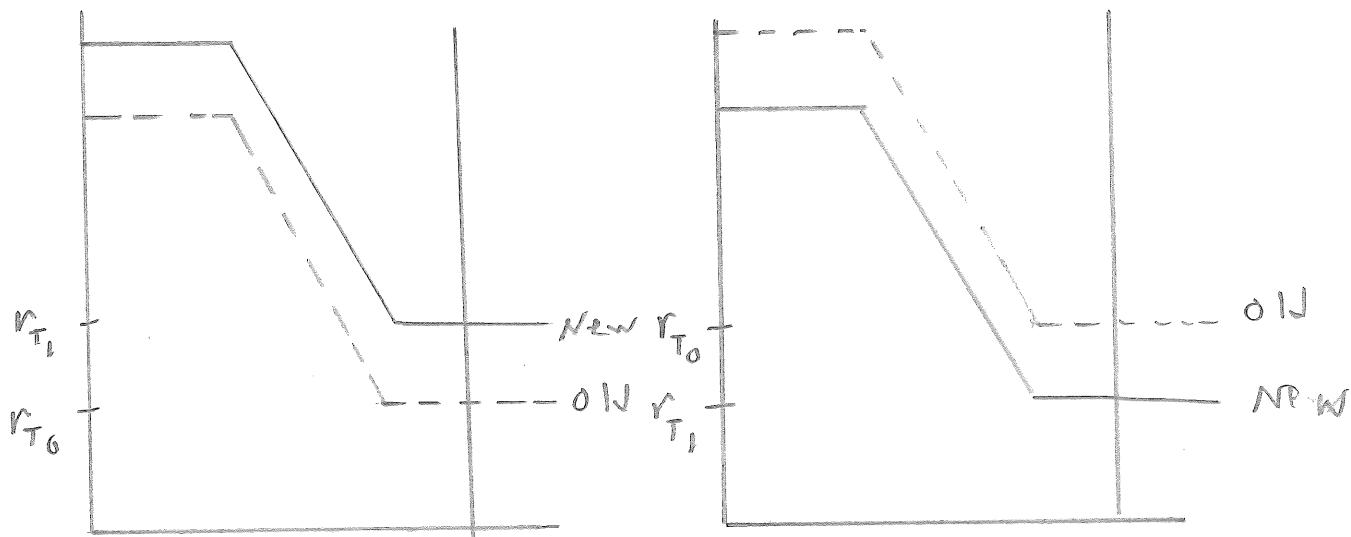
Note that there is no need for the central bank to change reserve supply. If something happens that would tend to reduce reserve supply a lot, so that the market rate might get off the floor, the central bank has to do open-market operations to keep reserve supply big enough to keep the market rate down on the floor. But this is not related to a change in r_T ; it's *always* true.

c) Corridor (or channel)

In the (symmetric) corridor system, the central bank sets r_P equal to r_T plus something, and r_D equal to r_T *minus* something. For example, it's rule might be that $r_P = r_T + 0.5\ percent$ and $r_D = r_T - 0.5\ percent$.

Or $r_P = r_T + 0.25\ percent$ and $r_D = r_T - 0.25\ percent$. Then it uses open-market operations to adjust reserve supply every day to keep the market rate equal to r_T:
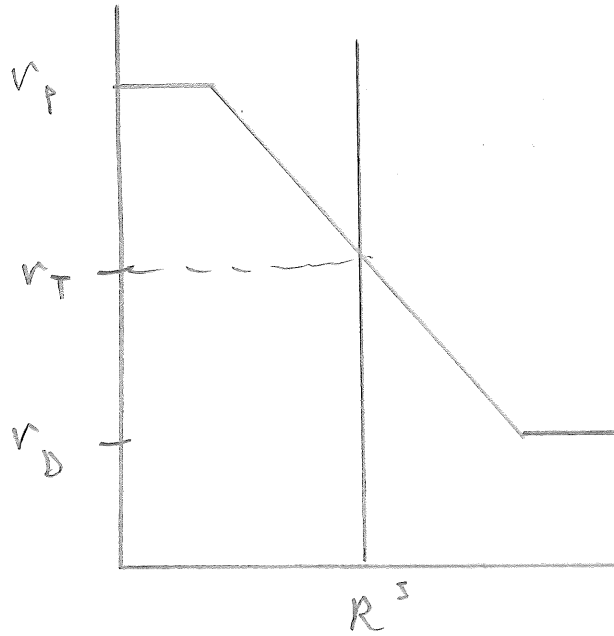
If the central bank doesn't get reserve supply exactly right on a day, the market rate won't be exactly equal to r_T on that day. But as long as the central bank gets it right on average the *expected* value of the market overnight rate will be equal to r_T. And that is all that matters for determination of interest rates in general. Note that, as in the floor system, the central bank need not change reserve supply when it changes r_T.

Increase in r_T Decrease in r_T
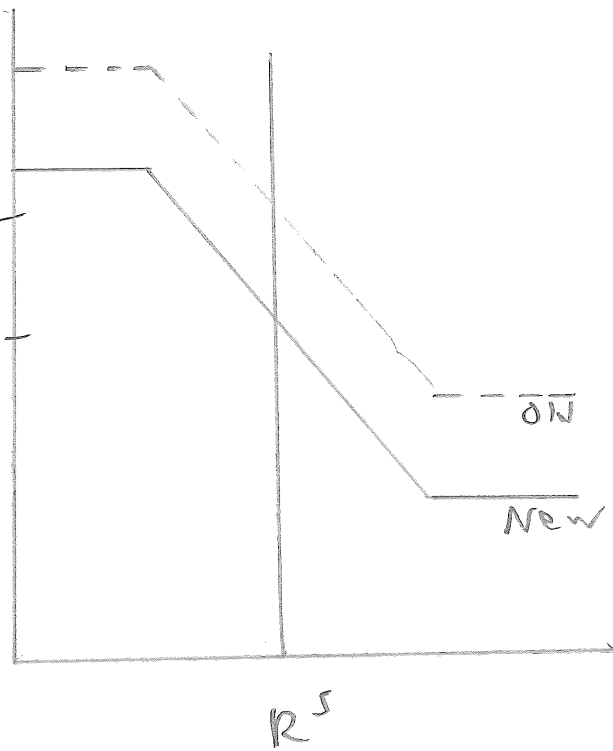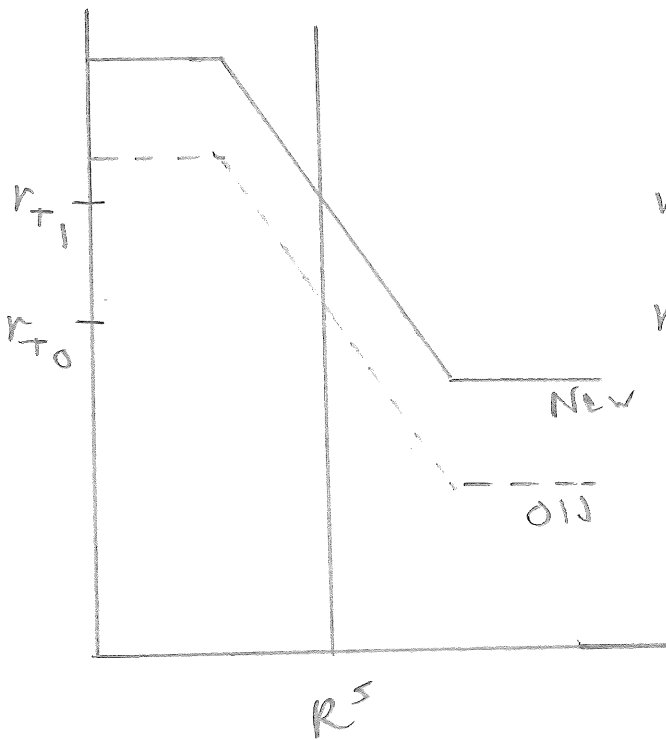
d) Fed's pre-2007 system

Prior to 2007, the Fed did not pay interest on reserves. In terms of our graphs so far, r_D was always zero. It did not always change r_P when it changed r_T. In fact, it usually held r_P at an unchanged value when it changed r_T. Thus, it usually had to *increase* reserve supply (buy stuff in open-market operations) when it cut r_T, and *decrease* reserve supply (sell stuff in open-market operations) when it hiked r_T.
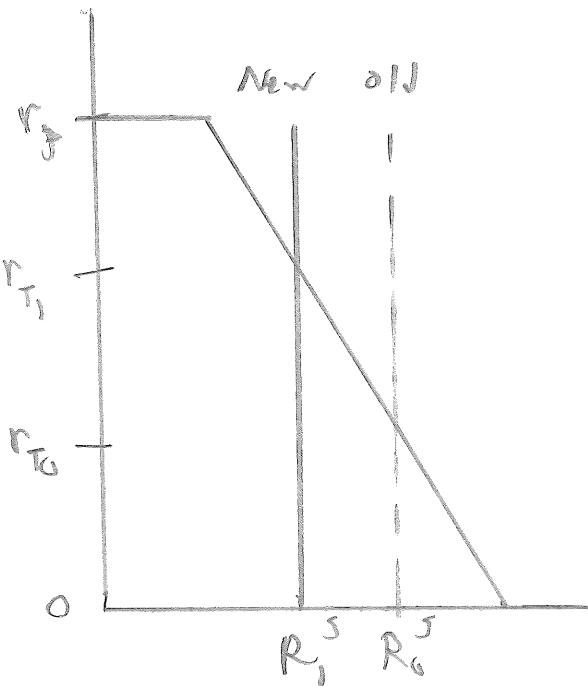
Increase in r_T Decrease in r_T
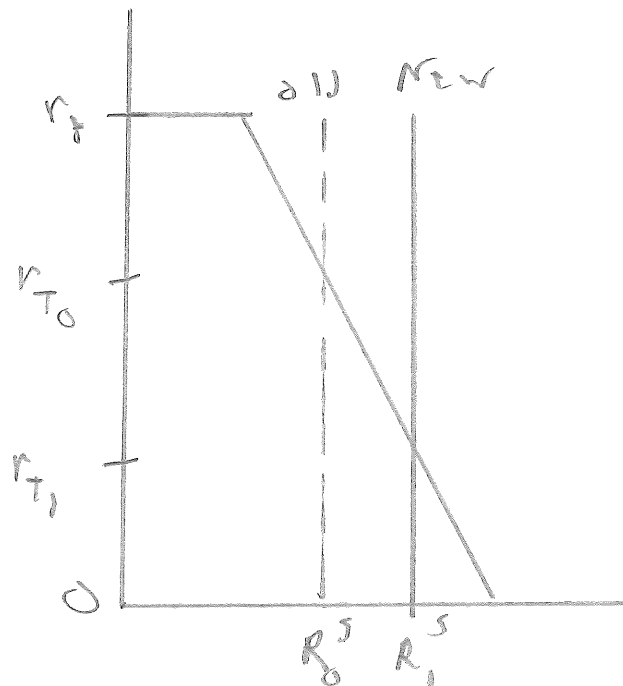
9) Complication in the U.S.: subfloor

a) In the U.S., *r* actually *can* fall below r_D

I said that the market overnight rate, which we have denoted by r, cannot fall below the rate that the central bank pays on reserve balances, r_D. But that's not true in the U.S. Our market overnight rate is actually a little bit lower than r_D. Why?

b) Some holders of reserve accounts don't get paid interest on reserves

Remember how I explained why the market rate can't fall below r_D: I said "just by leaving funds in its reserve account, a bank can lend to the Fed and get an interest rate of r_D. So no bank would lend to anyone else at a rate lower than r_D!"
That is true. But banks (and the U.S. Treasury) are not the only organizations that hold accounts in the Federal Reserve system. Organizations other than banks that hold Fed accounts include "Fannie Mae," "Freddie Mac," and "Federal Home Loan Banks" (FHLBs) (q.v.). Like banks, these organizations use their Fed accounts to make payments to banks and to the Federal government. But unlike banks, they *don't* get paid interest on their Fed accounts. Thus these other entities *are* willing to lend overnight at interest rates less than r_D.

Why doesn't the Fed pay interest to these other entities? Because the law that allowed the Fed to start paying interest on reserve accounts does not allow the Fed to pay interest to organizations other than banks.

This is a problem. To keep control over the sttructureureuremarket rate, the Fed can tolerate it if the market rate falls a *little* below r_D, but it wants to keep the market rate from falling much below r_D. So...

c) Solution: subfloor (ON RRP)

So the Fed has established a "subfloor" on the market overnight rate. It makes a standing offer to borrow overnight from anyone at an interest rate set a little below r_D. The Fed borrows through repo loans. When the fed borrows (rather than lends) through repo, it call it "reverse repo." So the rate it pays for these loans is the "Overnight reverse repo" (ONRRP) rate. *No one* needs to take less than the ONRRP rate for an overnight loan. So the market rate cannot fall below it.

E) Mathematical model of reserve demand

1) Introduction

I told you that reserve demand looked like:
I said that in between the flat parts,
reserve demand is negatively
related to the market overnight rate - the
demand curve is downward-sloping -
because the overnight rate is the
opportunity cost to a bank of holding a
dollar in its reserve account (rather than
lending it out overnight). I said that changes in
banks' required reserve balances shift
the demand curve left and right - increases
in required reserves shift the demand curve to the right;
decreases in required reserves shift the demand curve
to the left. I explained why the demand curve is flat at
the top and bottom. But there were several things I didn't explain.
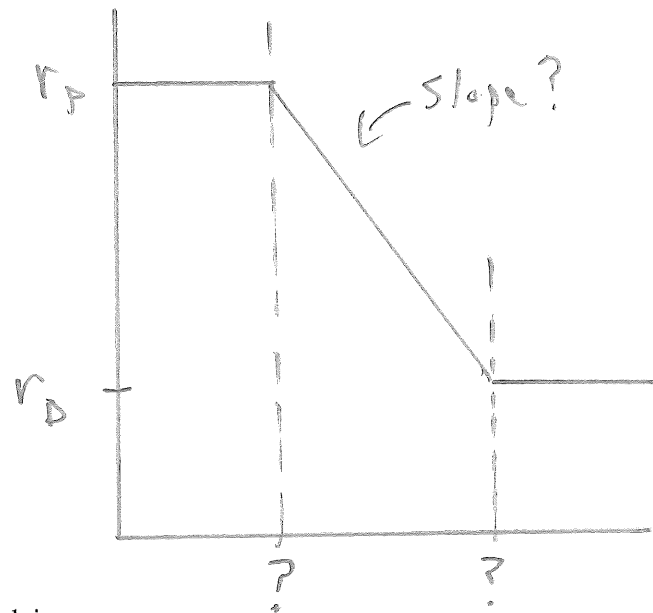
Why do changes in required reserve balances shift the reserve demand curve left and right?

Exactly why is the reserve-demand curve downward sloping between the flat parts? What is the *slope* of the reserve-demand curve between the flat parts?

And what about those corners marked *A* and *B* on the graph? Those corners correspond to specific quantities of reserves on the horizontal axis of the graph. What are those quantities?

In this section, I will show you a mathematical model that explains all those things. It relies on two ideas we have already used in this class. It also relies on some ideas you were supposed to learn in Stats: the c.d.f. (cumulative distrubition function) of a random variable, and the conditional expectation of a random variable. I will review all of these things before presenting the model.

The model is described in the coursepack reading *6) Understanding Monetary Policy Implementation.* I will use the same notation that the reading uses.

## 2) Review: things from this class

a) Liquidity preference

Think back to when I discussed demand for money in an economy with no banks and no central bank. I talked about the difference between *using* money and *holding* money. I said that a person who holds money loses something - the return to holding more bonds or making more loans - for every unit of money that he holds. Thus, we must explain why people would choose to hold money, rather than use money without holding it. I said that economists' explanation for this is "liquidity preference." This is the idea that even the most liquid interest-paying asset is a *little* bit illiquid. In the Baumol-Tobin model, the most liquid asset was bonds. The way bonds were illiquid was that there was a cost of buying or selling them (the cost F of going to the broker's office). The model showed that a person who wanted to maximize his "profit" - the return to his investments minus the cost of going to the broker - would hold a quantity of money that was negatively related to the return to holding bonds.

Here we are going to apply the idea of liquidity preference to a bank's demand for reserves. The "money" is reserves. The most liquid asset other than reserves is overnight loans. Overnight loans are a little illiquid. However, the way that overnight loans are illiquid is different from the way bonds are illiquid in the Baumol-Tobin model. You'll see what I mean below.

### b) Using calculus to find the optimal value of a choice variable.

In the Baumol-Tobin model, we needed to find the person's profit-maximizing value of N, where N was the number of financial transactions (visits to the broker). Once we knew N^*, we knew the person's average money balance, which was the quantity of money demanded. To find N^* we created an equation that gave profit Z (on the left-hand side of the equation) as a function of N and other stuff (F and Y). Then we used the math trick that relies on the fact that the slope of a line is zero at a maximum or minimum, and the slope is equal to the first derivative:
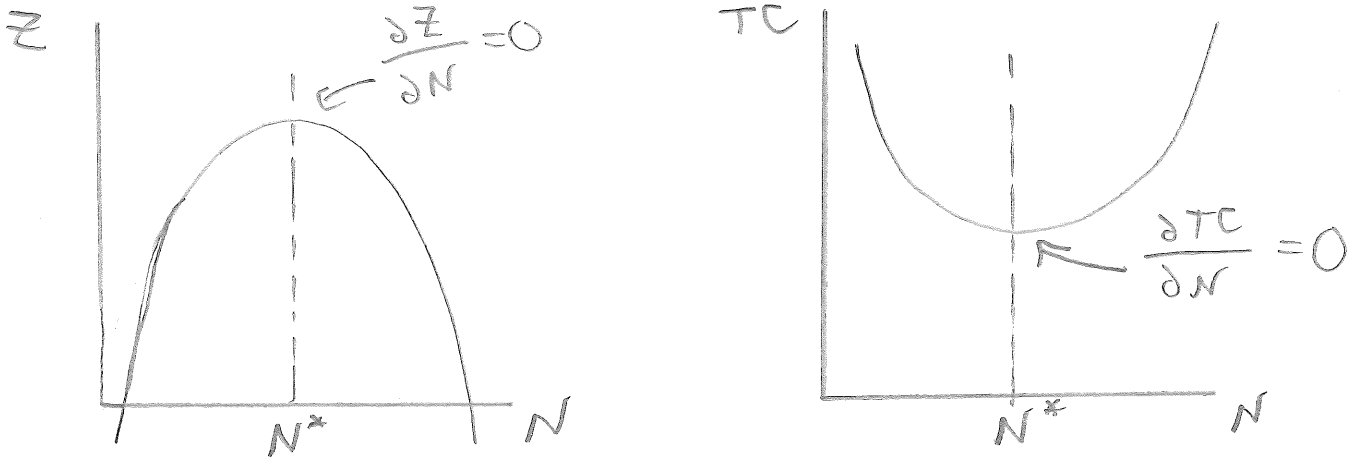
Thus, we took our equation that gives Z as a function of N, took the derivative of Z with respect to N and set it equal to zero. That gave another equation with zero on the left-hand side, and N on the right-hand side along with other stuff. We solved that equation for N to get N^*.

We will use the same trick here. We will create an equation that gives the bank's profit as a function of its reserve balance R and other stuff. The other stuff will include the market overnight rate r, the interest rate the central bank pays on reserves r_D, the interest rate the central bank charges for emergency loans r_p, and the bank's required reserve balance.

We'll take the derivative of the equation with respect to R and set it equal to zero. That will give another equation, with zero on the left-hand side and R on the right-hand side along with other stuff. Solving this equation for R gives R^*, which is reserve demand. The "quantity demanded" of reserves - the reserve balance a bank wants to hold - will be related to the market overnight rate r in a way that matches the reserve-demand curve we have been drawing.
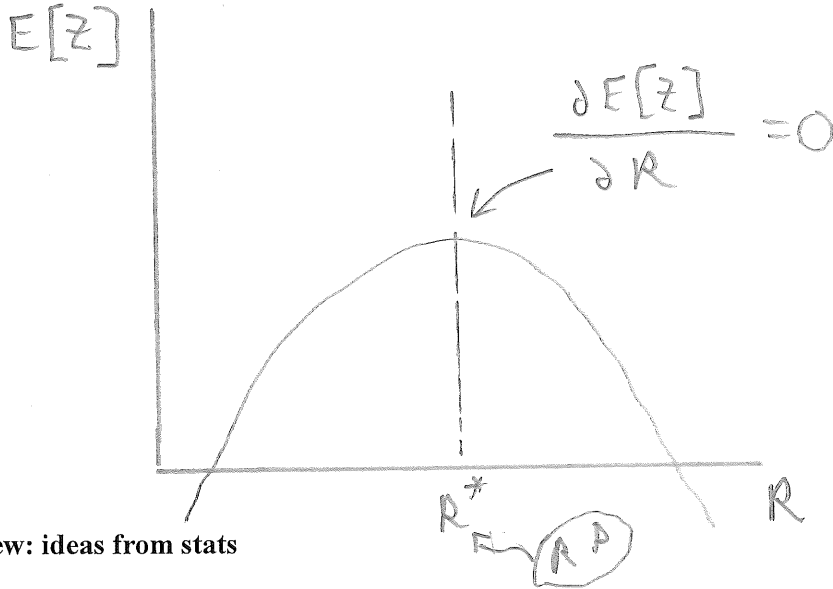

3) Review: ideas from stats

a) The distribution of a random variable and its histogram

Suppose we don't know the value a particular variable will take (or has taken - perhaps we just haven't observed it yet). But we know all the possible values it can take and the probability it will take each of those possible values. That is the *probability distribution* for the variable. You can describe a probability distribution with a table, or with a graph called a *histogram*. We used tables of probability distributions earlier when we talked about future values of the overnight rate in the context of the yield curve. For an example, say next year's overnight rate can take three possible values - 1.2, 1.5, 1.8 - and is equally likely to take any of those values. Then the probability distribution and histogram are:

Value	Probability
1.2	1/3
1.5	1/3
1.8	1/3
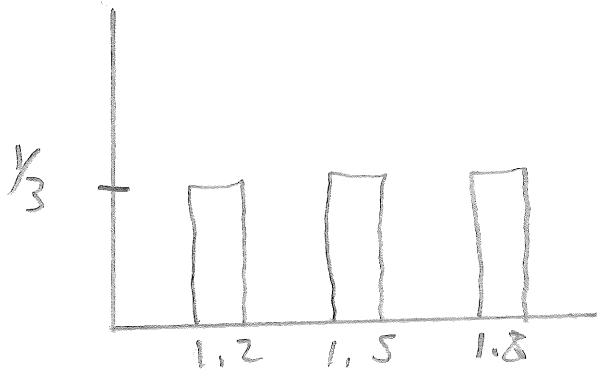
Another example is:

Value	Probability
1	1/6
2	1/6
3	2/6
4	1/6
5	1/6
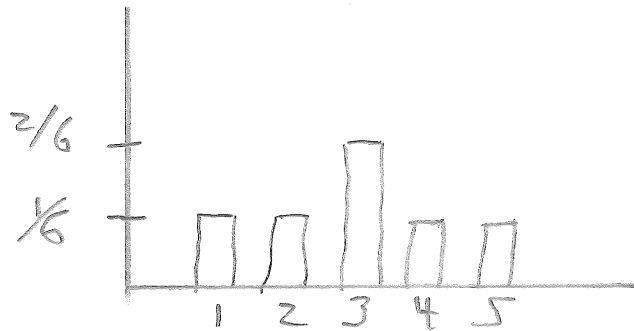
Of course, the probabilities of the possible values must add up to equal one.

b) The c.d.f. of a random variable

Yet another way to describe a distribution is with a function called a "cumulative distribution function" or c.d.f. A c.d.f. is denoted $F\{X\}$. You stick in a number for X and the c.d.f. gives you the probability that the random variable will turn out to be X or a smaller number. Here are the c.d.f.'s for the distributions I have used so far.

First example

Value	Probability	$F\{X\}$
1.2	1/3	1/3
1.5	1/3	2/3
1.8	1/3	3/3 = 1

$F\{1.2\} = 1/3$ because 1.2 is the lowest value the variable can take. So $F\{1.2\}$ is just the probability the variable will equal 1.2, which is 1/3. $F\{1.5\} = 2/3$ because the variable can equal 1.5 with probabilityy 1/3, or the lower number 1.2 with probability 1/3, so the total probability is 2/3. $F\{1.8\} = 1$ because the variable is certain to equal 1.8 or a lower value.

What is $F\{1.1\}$? Zero, because the variable cannot possibly turn out to be 1.1 or a smaller number. $F\{1.05\}$ is also zero, and so on. Generally, $F\{X\} = 0$ for any number that is less than the minimum possible value of the random variable.

What is $F\{1.9\}$? One, because the random variable is certain to take a value less than 1.9. Generally, $F\{X\} = 1$ for any number greater than (or equal to) the maximum possible value of the random variable.

Second example

Value	Probability	$F\{X\}$
1	1/6	1/6
2	1/6	2/6
3	2/6	4/6
4	1/6	5/6
5	1/6	6/9/ = 1

For any number less than 1, $F\{X\} = 0$. For any number greater than 5, $F\{X\} = 1$.

c) Conditional expectation (or expected value)

We have used expected values a lot. The expected value of a random variable P is denoted $E[P]$. As you know, you get it by multiplying the possible values times their probabilities, and adding up.

The conditional expectation is an expected value, but it is the expected value after you have some additional information about the random variable: you know the realized value will turn out to be within a certain range (or already is - perhaps you just haven't observed it yet). You know it is in that rangem but you don't know where in that range it is. To put it another way, you can *rule out* some of the possible values in the original probability distribution. is within a certain range, but you don't know what value it took within that range. The "conditional expected value" or "conditional expectation" of a variable is the answer to the question: what is the expected value of the variable given that you know the value is within that range, but that is all you know. It is the expected value of the variable on the *condition* that the realized value fell within that range. Frequently the range in question is an inequality, e.g. "I know the variable turned out to be less than 5," or "I know the variable turned out to be geater than or equal to negative 3."

A conditional expectation is denoted $E[P|condition]$. For example,

- if the condition is "I know the variable turned out to be less than 5," it's $E[P|P<5]$.

- if the condition is "I know the variable turned out to be greater than or equal to negative 3" it's $E[P|P \geq -3]$.

Here's some examples of expected values and conditional expectations from the distributions I have used so far.

First example

Value	Probability	$F\{X\}$
1.2	1/3	1/3
1.5	1/3	2/3
1.8	1/3	3/3 = 1

$$E[P] = \frac{1}{3} \times 1.2 + \frac{1}{3} \times 1.5 + \frac{1}{3} \times 1.8 = 1.5$$

$E[P \mid P \leq 1.2] = 1.2$ (because the only possible value of P on this condition is 1.2)

$$E[P \mid P \leq 1.5] = \frac{1}{1/3 + 1/3}\left(\frac{1}{3} \times 1.2 + \frac{1}{3} \times 1.5\right) = 1.35$$

Second example

Value	Probability	$F\{X\}$
1	1/6	1/6
2	1/6	2/6
3	2/6	4/6
4	1/6	5/6
5	1/6	6/9/ = 1

$$E[P] = \frac{1}{6} \times 1 + \frac{1}{6} \times 2 + \frac{2}{6} \times 3 + \frac{1}{6} \times 4 + \frac{1}{6} \times 5 = 3$$

$$E[P \mid P \leq 2] = \frac{1}{1/6 + 1/6}\left[\frac{1}{6} \times 1 + \frac{1}{6} \times 2\right] = 1\tfrac{1}{2}$$

$$E[P \mid P \leq 3] = \frac{1}{1/6 + 1/6 + 2/6}\left[\frac{1}{6} \times 1 + \frac{1}{6} \times 2 + \frac{2}{6} \times 3\right] = 2\tfrac{1}{4}$$

d) The uniform distribution

i) Definition

In Stats class you learned about some standard distributions that seem to describe many random variables in reality. Two of these are the *normal* distribution and the *Poisson* distribution. The histograms for these distributions look like:

Normal distribution   Poisson distribution:
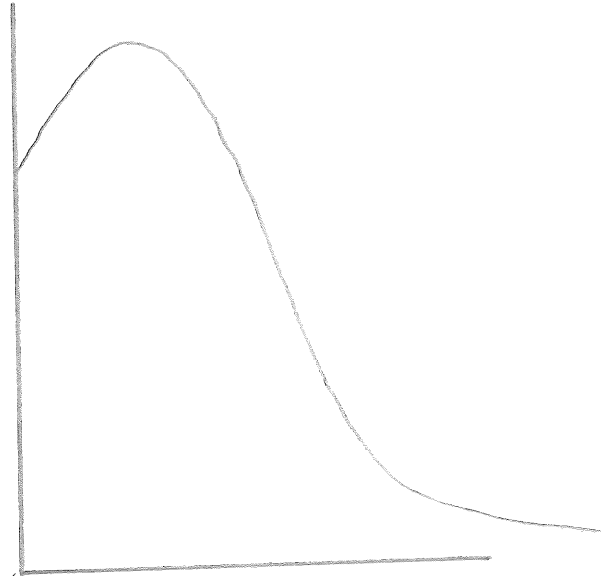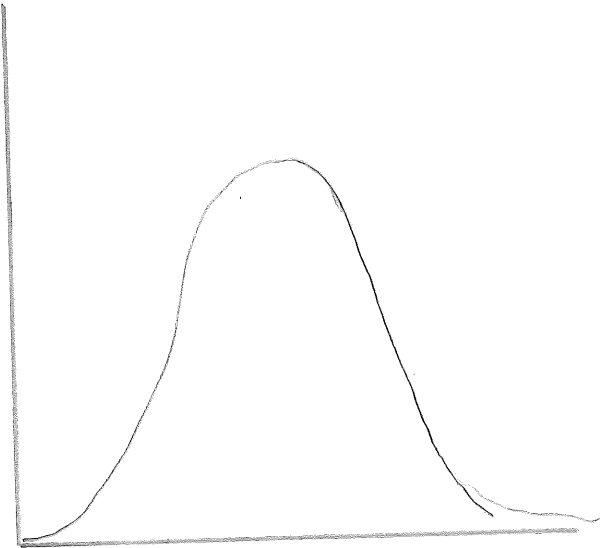
A good thing about these standard distributions is that you don't need to make tables to figure out their expected values, c.d.f.s and conditional expected values. There are mathematical formulas you can use to calculate all those things. Unfortunately, the formulas are complicated.

There is one standard distribution for which the formulas are easy. It is the "uniform distribution." A variable that follows a uniform distribution has a maximum value and a minimum value, and it is equally likely to take any value within those bounds. An example of a uniform distribution is:

Value	Probability	$F\{X\}$
7	1/11	1/11
8	1/11	2/11
9	1/11	3/11
10	1/11	4/11
11	1/11	5/11
12	1/11	6/11
13	1/11	7/11
14	1/11	8/11
15	1/11	9/11
16	1/11	10/11
17	1/11	11/11 = 1

$$E[P] = \frac{1}{11} \times 7 + \frac{1}{11} \times 8 + \frac{1}{11} \times 9 + \dots \frac{1}{11} \times 17 = 12$$
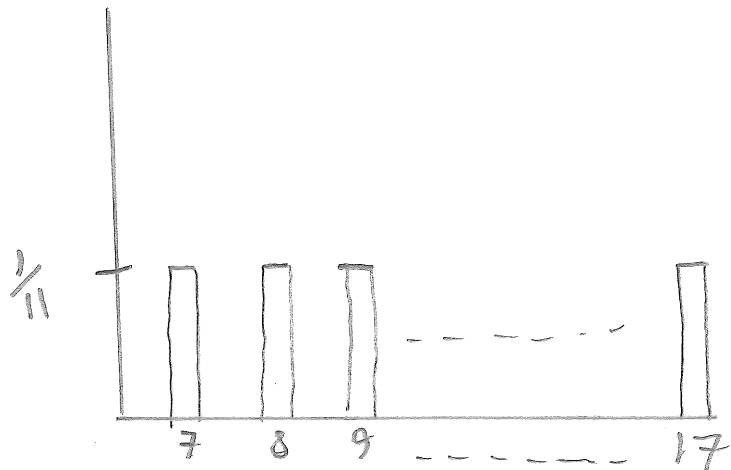
I think of a uniiform distribution as a line. The left endpoint of the line is the minimum value. The right endpoint is the maximum value. Like:
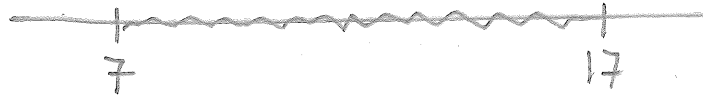
Someone is throwing darts at the line. A dart must land somewhere on the line, but it is equally likely to hit any point on the line. Where the dart hits, that's the realized value of the random variable.

ii) Formulas

Let P be a uniformly distrubuted random variable. P_{Min} is its minimum possible value. P_{Max} is its maximum possible value.

Expected value

The expected value of P is $E[P] = \frac{1}{2}(P_{Min} + P_{Max})$. The expected value is just the average of the minimum and maximum values.

C.d.f.

The c.d.f, that is the probability that P will turn out to be less than or equal to X, is:

- $F\{X\} = \dfrac{X - P_{Min}}{P_{Max} - P_{Min}}$ for values of X between P_{Min} and P_{Max}

- $F\{X\} = 0$ for values of X less than P_{Min}

- $F\{X\} = 1$ for values of X greater than P_{Max}

Conditional expected values

Say you know that P turned out to be less than or equal to a number X on the line, that is the random variable hit somewhere on the segment of the line between P_{Min} and X, but that is *all* you know. Then:

$E[P \mid P \leq X] = \frac{1}{2}(P_{Min} + X)$ (Like a uniform distribution with a minimum P_{Min} and a maximum X.)

Say you know that P turned out to be greater than or equal to a number X on the line, that is the random variable hit somewhere on the segment of the line between X and P_{Max}, but that is *all* you know. Then:

$E[P \mid P \geq X] = \frac{1}{2}(X + P_{Max})$ (Like a uniform distribution with a minimum X and a maximum P_{Max}.)

4) Liquidity preference in the context of reserve demand

The way that overnight loans are a slightly-illiquid asset is a result of the way the Fed settles payments between reserve accounts.

The Fed is linked to banks with a communications network called "Fedwire." Each day, the Fed takes payment orders from banks through Fedwire. The orders say things like "Take $1.5 million out of my reserve account and transfer it to Citibank's reserve account." Banks use Fedwire orders to clear checks and other payments between people with accounts at different banks. Banks also use Fedwire orders to make or receive overnight loans. If Bank A lends $100 million overnight to Bank B, Bank A sends a Fedwire message to the Fed that says "take $100 million out of my reserve account and transfer it to the reserve account of Bank B." If Bank B borrows $100 million overnight from a *customer* of Bank A, Bank A sends the same message to Fedwire (and subtracts $100 million from the lender's checking account).

At a certain time each day, say 6 pm, the Fed shuts off the Fedwire network. At that time, it becomes impossible for a bank to make or take any more overnight loans.

But the Fed keeps processing Fedwire orders for a while after 6 pm. For about an hour, it continues to debit funds from banks that ordered payments and crediting funds to bank receiving payments.

At 7 pm the Fed is finished processing. There is a final balance in each bank's reserve account. The balance will probably be different from the balance that the bank left in the account at 6 pm. If there were a lot of debits to the account between 6 and 7, the final balance will be less than the 6 o'clock balance. If there were a lot of credits to the account between 6 and 7, the final balance will be more than the 6 o'clock balance.

A bank has an overdraft in its account if the final balance is less than zero. To cover the overdraft, the band must borrow from the Fed at the rate r_p.

If it is the last day of the "maintenance period" (see part C) 6) above), the Fed compares the final balance with the amount that the bank must hold in the account to satisfy its reserve requirement. If the final balance is not enough, the bank must borrow from the Fed at r_p to bring the balance up to the required level.

Importantly, as of 6 pm a bank cannot perfectly predict what its reserve balance will be at 7 pm, because it cannot be sure which payment orders will be processed on a given day. Some payments are processed with an unpredictable lag - might be processed today or tomorrow, can't be sure which. Thus, the bank does not know exactly what its final balance for the day will be. That means the bank may have a shortfall in its reserve account (an overdraft or failure to meet its reserve requirement) even if it was *aiming* to have a zero balance (to avoid overdraft) or positive balance (to satisfy its reserve requirement).

A bank finds itself with a reserve shortfall at 7 pm would like to withdraw from some of its overnight loans or borrow some money overnight to get cash to cover the shortfall, so that it doesn't have to borrow from the Fed at the primary credit rate r_p. But it can't - it's too late. That is the way overnight loans are a little illiquid, as far as a bank is concerned.

5) The model

a) Story and notation

The model describes one bank's reserve demand. To get total reserve demand, we say all banks are identical and just like the bank described by the model.

For simplicity, we say that there is no multi-day maintenance period - a bank must satisfy its reserve requirement on every day. Also for simplicity we say that the Fed pays interest only on a bank's *excess* reserve balance, that is the bank's

reserve balance *minus* the required reserve balance. In reality the Fed pays interest on a bank's entire reserve balance. But this makes the notation more complicated and it doesn't affect the answer.

Let A denote the total amount of money that the bank has available just before 6 pm, which it can choose to lend overnight or leave in its reserve account. Then the amount that the bank lends overnight is A-R. The interest it will earn on its overnight lending is *r (A-R)*.

At 7 pm the Fed is finished settling payments. There is a final balance in each bank's reserve account.

The final balance is equal to R if the credits to the reserve account (incoming payments) processed after 6 were just equal to the debits (outgoing paymments).

The final balance is smaller than R if the debits were bigger than the credits.

The final balance is greater than R is the debits were bigger than the credits.

P denotes the net credit or debit. Thus the amount left in the account at 7 pm is $R+P$, where P can be positive (credits bigger than debits), negative (debits bigger than credits), or zero.

The Fed compares the final balance $R+P$ with the bank's required reserve balance. K is the required reserve balance.

If $R+P > K$, the bank has satisfied its reserve requirement. It has an excess reserve equal to $R+P$-K . It will be paid interest on its excess reserve equal to $r_D \times (R + P - K)$.

If $R+P < K$, the bank has a shortfall equal to K - *(R+P)*, which it must borrow from the Fed. It will pay interest on the loan equal to $r_p \times (K - R - P)$.

At 6 pm the bank does not know exactly what P will turn out to be.

b) What a bank wants to maximize

The bank wants to maximize profit, which we denote Z. (as we did for the Baumol-Tobin model). Z will be equal to:

the interest on the bank's overnight lending *r (A-R)*
minus any interest it has to pay the Fed on a primary credit loan to cover a reserve shortfall $r_p \times (K - R - P)$
plus any interest it is paid by the Fed on excess reserves $r_D \times (R + P - K)$.

But at the time it must choose R, however, that is at 6 pm, the bank may not know exactly what profit will result from any given value of R it might choose, because it does not know exactly what P will be. The bank may not know for sure whether it will end up with excess reserves or a reserve shortfall. It may not know for sure how big the shortfall will be *if* it runs a shortfall, or how big the excess will be *if* it has excess.

Thus, we say that the bank chooses R to maximize the *expected value* of its profit *E[Z]*, which is:

The interest on the bank's overnight lending *r (A-R)*

minus the probability the bank has a reserve shortfall times the expected value of its interest payment to the Fed *assuming* the bank has a reserve shortfall

plus the probability the bank ends up with excess reserves *times* the expected value of its interest earnings *assuming* the bank has excess reserves.

24

Notice that these expected values are *conditional* expectations. The conditions are "*assuming* the bank has excess reserves" and "*assuming* the bank has a reserve shortfall." Using the notation for conditional expected values, we can say:

$$E[Z] = r\,(A\text{-}R)$$
$$-\ (Prob.\ R\text{+}P < K) \times r_P \times E[K - R - P \mid R + P > K]$$
$$+\ (Prob.\ R\text{+}P \geq K)\ \times r_D \times E[R + P - K \mid R + P > K]$$

To figure out those probabilities and condition expected values, we need a probability distribution for P. This is where the uniform distribution comes in.

c) Uniform distribution for P

Say the bank's probability distribution for P is uniform. The minimum possible value of P is $-\overline{P}$. The maximum possible value of P is $+\overline{P}$. So:

$$E[P] = \frac{1}{2}\left(-\overline{P} + \overline{P}\right) = 0$$

$$F\{X\} = \frac{X - (-\overline{P})}{\overline{P} - (-\overline{P})} = \frac{X + \overline{P}}{\overline{P} + \overline{P}} = \frac{X + \overline{P}}{2\overline{P}} = \frac{\overline{P}}{2\overline{P}} + \frac{1}{2\overline{P}}X = \frac{1}{2} + \frac{1}{2\overline{P}}X$$

$$E[P \mid P \leq X] = \frac{1}{2}\left(-\overline{P} + X\right)$$

$$E[P \mid P \geq X] = \frac{1}{2}\left(X + \overline{P}\right)$$

Because P has a minimum possible value (it cannot be lower than $-\overline{P}$) it is possible for a bank to make sure it does not suffer a reserve shortfall at 7 pm. It can do this by leaving a large enough balance R in its reserve account at 6 pm. A balance greater than or equal to $R + \overline{P}$ will do the trick. The most that can be subtracted from the bank's account between 6 and 7 pm is $\overline{P}$. So if $R \geq K + \overline{P}$ then even if $P = -\overline{P}$ the final balance $R\text{+}P$ must be greater than or equal to K.

It is also possible for a bank to make sure it *does* suffer a reserve shortfall at 7 pm, because P has a maximum possible value (it cannot be greater than $+\overline{P}$). It can do this by leaving a *small* enough balance in the account at 6 pm. A balance less than $R - \overline{P}$ will do the trick. The most that can be added to the bank's account between 6 and 7 pm is $\overline{P}$. If $R < K - \overline{P}$, then even if $P = +\overline{P}$ the final balance $R\text{+}P$ must be less than K.

What if R is more than $K - \overline{P}$ but less than $K + \overline{P}$? Then the bank may or may not have a shortfall, depending on what P turns out to be. Using the c.d.f for the uniform distribution, we can figure out the *probability* of a shortfall.

A shortfall occurs if $R + P < K$, which is to say a shortfall occurs if $P < K - R$. The probability that $P < K - R$ is $F\{K - R\}$. So the probability of a shortfall is:

$$F\{K - R\} = \frac{1}{2} + \frac{1}{2\overline{P}}(K - R)$$

The probability that a shortfall does *not* occur, which is to say the probability there will be excess in the account, is:

$$1 - F\{K - R\} = 1 - \left(\frac{1}{2} + \frac{1}{2\overline{P}}(K - R)\right) = \frac{1}{2} - \frac{1}{2\overline{P}}(K - R)$$

We can figure out the expected value of a shortfall, assuming a shortfall occurs, using the uniform distribution's formulas for conditional expected values. The amount of a shortfall is $K - R - P$.

The expected value of a shortfall is $E[K - R - P]$.

K and R are known values as of 6 pm, so $E[K - R - P] = K - R - E[P]$.

"Assuming a shortfall occurs" is the condition that $P < K - R$. So the expected value of a shortfall assuming that a shortfall occurs is:

$$K - R - E[P \mid P < K - R] = K - R - \frac{1}{2}\left(-\overline{P} + K - R\right) = K - R + \frac{1}{2}\overline{P} - \frac{1}{2}K + \frac{1}{2}R = \frac{1}{2}\overline{P} + \frac{1}{2}K - \frac{1}{2}R = \frac{1}{2}\left(\overline{P} + K - R\right)$$

Finally, we can figure out the expected value of excess assuming there is excess. The amount of excess is $R + P - K$.

The expected value of a shortfall is $E[K - R - P]$.

K and R are known values as of 6 pm, so $E[R + P - K] = R - K + E[P]$.

"Assuming there is excess" is the condition that $P > K - R$. So the expected value of excess assuming that there is excess is:

$$R - K + E[P \mid P \geq K - R] = R - K + \frac{1}{2}\left(K - R + \overline{P}\right) = \frac{1}{2}R - \frac{1}{2}K + \frac{1}{2}\overline{P} = \frac{1}{2}\left(\overline{P} - K + R\right)$$

d) Reserve demand

Now we can figure out why the things we wanted to know about the reserve demand curve.

First, what reserve balance R will a bank want to hold if $r = r_D$?

It will want to hold a reserve balance large enough to ensure that it will *not* run a reserve shortfall, which is any balance greater than or equal to $K + \overline{P}$. Why?

Because when $r = r_D$ there is no cost to holding dollars in the reserve account - the bank is paid as much interest on a dollar in the reserve account as it can earn lending that dollar out overnight. The bank might as well hold enough in the reserve account to entirely eliminate the danger of a reserve shortfall, because running a shortfall is costly: the bank will have to pay the high interest rate r_P to cover it. Thus, we can say that when $r = r_D$, $R^D \geq K + \overline{P}$.
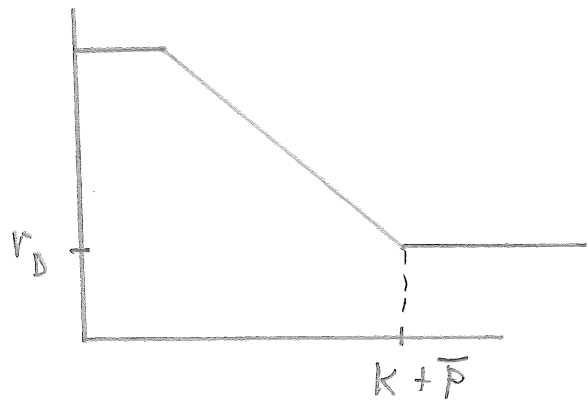

Second, what if $r = r_P$? Then the bank will want to hold a reserve balance small enough to ensure that it will run a reserve shortfall, which is any balance less than $K - \overline{P}$. Why?

Because running a shortfall is costless: the market overnight rate, which is the gain from holding less in the reserve account (the bank can lend more overnight) is equal to the rate the bank pays to borrow a dollar from the Fed by running an overdraft. The bank might as well make sure that it will not be left with any excess in the account, because excess is costly: on every dollar of excess reserve the bank will earn only the low interest rate r_D. Thus, we can say that when $r = r_P$, $R^D < K + \overline{P}$.

26
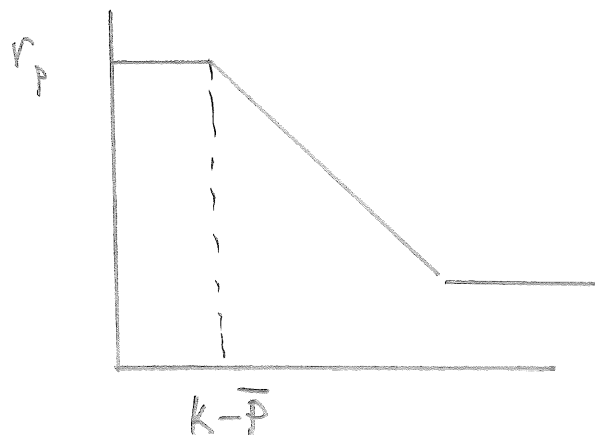
Finally, what if r is between r_p and r_D? Then the bank will choose the value of R that maximizes:

$$E[Z] = r\ (A\text{-}R)$$
$$-\ (\text{Prob. } R+P < K) \times r_p \times E[K - R - P \mid R+P < K]$$
$$+\ (\text{Prob. } R+P > K)\ \times r_D \times E[R+P-K \mid R+P > K]$$

which is:

$$E[Z] = r\ (A\text{-}R)$$
$$-\left[\frac{1}{2} + \frac{1}{2\overline{P}}(K-R)\right] \times r^P \times \left[\frac{1}{2}\left(\overline{P}+K-R\right)\right]$$
$$+\left[\frac{1}{2} - \frac{1}{2\overline{P}}(K-R)\right] \times r_D \times \left[\frac{1}{2}\left(\overline{P}-K+R\right)\right]$$

See that this is a big equation with R in it! So now we can use our calculus trick. Take the derivative of this equation with respect to R and set it equal to zero:

$$0 = \frac{\partial E[Z]}{\partial R} = \ldots\ldots\ldots\ldots$$

That gives another equation, with zero on the left-hand side and R on the right-hand side along with other stuff. Solving this equation for R gives $R*$, which is reserve demand R^D when r is between r_p and r_D. The answer that you get should be:

$$R_D = \quad K + \overline{P}\left(\frac{r_P - 2r + r_D}{r_P - r_D}\right) \quad = \quad K + \overline{P}\left(\frac{r_P + r_D}{r_P - r_D}\right) \quad -\left(\frac{2\overline{P}}{r_P - r_D}\right)r$$

Notice that the coefficient on r is negative. That's good! It means reserve demand is negatively related to the market overnight rate. (Remember that we are not talking about the flat parts of the reserve demand curve. We are talking about reserve demand when r is between r_p and r_D.)

Notice also that, for any given value of r, increases in K increase reserve demand. Decreases in K have the opposite effect. Also good! It means increases in K shift the reserve demand curve to the right, and decreases in K shift it to the left.

What about the slope of the reserve demand curve? You can see that in the equation but you have to be careful. The reserve demand curve has r on the vertical axis, R on the horizontal axis. The equation we have derived gives the quantity of reserves demanded on the left-hand side, r (and other stuff) on the right-hand side. So you have to "invert" this equation to get the equation that, when you plot it, is the reserve demand curve. By "invert" I mean do algebra to get r on the left-hand side of the equation and R on the right-hand side. When you do that you get:

$$r = (r_P - r_D)\left(\frac{1}{2} + \frac{K}{2\overline{P}}\right) - \left(\frac{r_P - r_D}{2\overline{P}}\right)R_D$$

The coefficient on R_D in this inverted equation is the slope of the reserve demand curve: $-\left(\dfrac{r_P - r_D}{2\overline{P}}\right)$.

But I bet you are thinking "what about that derivative and algebra business? That must be a huge mess. I can't deal with that!" Don't worry. I don't expect you to. What I expect you to do is deal with a simplified version of the problem. For the

simplified version I say the Fed is not paying interest on reserves ($r_D = 0$) and there is no reserve requirement ($K = 0$). In that case,

$$E[Z] = r\,(A\text{-}R) - \left[\frac{1}{2} + \frac{1}{2\overline{P}}(-R)\right] \bullet r^P \bullet \left[\frac{1}{2}\overline{P} - \frac{1}{2}R\right]$$

Also, I give you an actual number to use for $\overline{P}$.

XIV) Monetary policy and macroeconomics, simplified
    A) Introduction
        1) What this section is about
        2) Overview
        3) Simplifications for this section
    B) The inflation target
        1) What the Federal Reserve Act requires the Fed to do
        2) What the Fed actually does
        3) Why 2 percent?
    C) The Phillips curve
        1) Introduction
        2) Empirical Phillips curves
        3) The expectations-augmented Phillips curve
            a) Review: variables and parameters
            b) Functions expressing the expectations-augmented Phillips curve
            c) The expectations-augmented Phillips curve with the unemployment rate u
            d) The expectations-augmented Phillips curve with real GDP Y
            e) Trend growth in Y^*
        4) Controlling inflation through the Phillips curve
    D) The IS curve
        1) Review and introduction
        2) The natural rate of interest
        3) Controlling inflation through the real interest rate
    E) Review: how to be an inflation-targeting central banker
    F) More about expected future inflation $E\pi$
        1) Introduction
        2) Anchored versus adaptive expectations
        3) History of $E\pi$ in America
        4) What central bankers want $E\pi$ to be
    G) Monetary policy when $E\pi = \pi^T$
        1) Introduction
        2) Changes in r^* due to shifts in the IS curve
            a) On a graph
            b) Fiscal policy
            c) Financial crises, stock market booms and crashes
        3) Changes in r^* due to changes in natural rates
            a) Changes in u^*
            b) Changes in Y^* due to fluctuations in productivity growth
        4) Accidental recessions and booms
            a) Fed officials can be wrong about r^*
            b) Accidental recessions and booms and deviations of inflation from target
            c) Examples
                i) Fed economists are wrong about the position of the IS curve
                ii) Fed economists are wrong about u^*
    H) The lower bound problem
        1) i^*
        2) Sometimes r^* is negative
        3) So i^* can be negative
        4) The lower bound on the nominal interest rate
            a) Introduction
            b) The zero lower bound

c) The effective lower bound

d) The lower bound can prevent a central bank from keeping $r=r*$

5) Result: recession

6) A danger: expectations become unanchored

7) An obvious fix: fiscal policy. But...

8) Solutions to the lower bound problem: changing the inflatin target

a) Introduction

b) Raise the inflaiton target

c) Switch from inflation targeting to price level targeting

XIV) Monetary policy and macroeconomics, simplified

A) Introduction

1) What this section is about

The previous section was about how the Fed keeps the market overnight rate equal to the target chosen by the FOMC. *This* section is about how the FOMC chooses the target overnight rate. Why would it set the overnight-rate target at 1 percent, or 2 percent? When it raises or lowers the target, why does it change the target in that direction and that much?

2) Overview

Recall that all interest rates in the economy, including bond yields, are ultimately based on the *expected value of the average value of the FOMC's target overnight rate* between now and the time the bond or loan will be paid off:

$$_n i_t = E\left[\frac{i^T_d + i^T_{d+1} + \cdots}{number\ of\ days}\right] + {}_n\tau + other\ premiums$$

The FOMC sets a target overnight rate, not because it cares about the overnight rate *per se*, but because it is trying to influence the general level of interest rates on bonds and loans. Sometimes the FOMC wants to keep that general level stable; sometimes the FOMC wants to raise it; sometimes the FOMC wants to lower it.

The FOMC wants to control the general level of interest rates because it believes that the interest rate level affects total spending on American-produced goods and services - aggregate demand - hence real GDP, which we denote by Y. Other things equal, an increase (decrease) in interest rates tends to reduce (increase) Y. What I am talking about here is the IS curve from Econ 362. Real GDP Y is relayted to the unemployment rate, denoted u. Other things equal, an increase (decrease) in Y is associated with a decrease (increase) in u. (The exact quantitative relationship between changes in Y and changes in u is called "Okun's Law.) There is a phrase that refers to *both* real GDP and the unemployment rate, together: "real activity." "An increase in real activity" means Y up, u down. "A decrease in real activity" means Y down, u up.

Real activity is related to inflation, which we denote by π . Other things equal, an increase in real activity (Y up, u down) is associated with an increase in inflation; a decrease in real activity (Y down, u up) is associated with a decrease in inflation. What I am talking about here is the Phillips curve from Econ 362.

The Fed has a target inflation rate. (Don't confuse the target inflation rate with the target overnight rate!) The Fed's target inflation rate is 2 percent. The Fed tries to keep inflation equal to 2 percent and, *subject to that*, keep unemployment as low as possible.

So this is what the Fed is up to with the target overnight rate. It tries to set the target overnight rate at the value that will create expectations of *future* target rates that will hold the general level of interest rates at the value that will keep real activity at the value that will keep inflation at two percent. Read that sentence again.

3) Simplifications for this section

For this section, to make the basic story clear, I will make three simplifications. First, I will assume that real activity is the *only* thing that affects inflation. (This isn't true; inflation is also affected by other factors that we call "supply shocks" or "cost-push shocks."). Second, I will assume that the interest rate that affects real activity is the overnight rate, that is the rate that the Fed can control directly using the techniques described in the previous section. (Actually, it is the general level of *longer-term* interest rates that affects real activity.) Finally, I will assume that changes in the real interest rate have an *immediate* effect on real activity and inflation. (Actually, there is a lag in this effect.) I will return to all of these points in the following section.

A) The inflation target

1) What the Federal Reserve Act requires the Fed to do

The law governing the Federal Reserve, the Federal Reserve Act, requires the Fed to *"promote effectively the goals of maximum employment, stable prices, and moderate long term interest rates."*

2) What the Fed actually does

Over the 1980s and early 1990s Federal Reserve policymakers came to a consensus that they would define "stable prices" to mean 2 percent inflation. There are many price indexes which can give different numbers for inflation, so Fed policymakers also defined which price index they have in mind: the PCE (personal consumption expenditure) price index. This is similar to the Consumer Price Index, but more reliable. For many years Fed policymakers did not *officially announce* that they had this inflation target, but they hinted at it very strongly. Finally, in 2012 they made it official, and they said that their inflation target was a "symmetric" objective. One meaning of the word "symmetric" is "balanced on both sides." When Fed people said their inflation target was symmetric objective, they meant that from their point of view a one percent miss on the high side is just as bad as a one percent miss on the low side - no better, no worse.

But "stable prices" is just one of the goals Congress set for the Fed. What about "maximum employment" and "moderate long term interest rates"? Fed policymakers pretty much ignore "moderate long term interest rates." What does that even mean? They don't ignore "maximum employment," but they interpret it to mean they should keep the unemployment rate as low as they can *as long as* inflation does not go higher than 2 percent. They talk about their "dual mandate" in the Federal Reserve Act, meaning the two things "stable prices" and "maximum employment." Congress and the President have always been OK with this, even though it doesn't seem entirely consistent with the text of the law, perhaps because everyone knows that the text of the law really makes no sense.

3) Why 2 percent?

Nearly all central banks have inflation targets and nearly all inflation targets are 2 percent. This is because, when inflation is *higher* than 2 percent, people notice that inflation is happening and they *really* don't like it. (We learned this during America's years of high inflation in the 1970s-1980s.) Economists don't understand why people hate noticeable inflation so much, but they do, so there. Why don't central banks target zero inflation? Mainly because of the lower bound problem, which I will discuss later in this section.

B) The Phillips curve

1) Introduction

The phrase "Phillips curve" refers to the relationship between inflation and real activity. But people use the phrase to refer to two different types of relationship. One is the "empirical" relationship. "Empirical" means "originating in or based on observation or experience." An "empirical" Phillips curve is the relationship that you can observe in statistics on inflation and real activity (the unemployment rate or real GDP) by looking at tables of the numbers, or by making scatterplots, or by running regressions. The other is a theoretical relationship that one cannot observe directly, but that economists believe

to be the *underlying mechanism* that creates the empirical relationships we see. This theoretical relationship is called the "expectations-augmented Phillips curve." In class, I will explain how the expectations-augmented Phillips curve could create the empirical Phillips curves we observe.

2) Empirical Phillips curves

The empirical relationship is, generally, that an increase in real activity from one year to the next (increase in Y, decrease in u) is usually accompanied by an increase in inflation. The exact relationship is a little different depending on what historical era you are looking at.

In most eras - 1870-1914, 1950-1967, late 1990s through today - what you see is a simple positive relationship between real activity and inflation. In years when real activity is especially high (that is, real GDP is high, unemployment is low), inflation tends to be especially high too; when real activity is low so is inflation.

In the era from the late 1960s to the late 1990s, what you see is a positive relationship between real activity and the *change* in inflation: in years when real activity was high inflation was accelerating (speeding up); when real activity was low inflation was decelerating (slowing down). The 1960s-1990s pattern is called the "accelerationist" Phillips curve.

3) The expectations-augmented Phillips curve

a) Review: variables and parameters

We describe the expectations-augmented Phillips curve with mathematical functions. Before we do that, let's review "variables" versus "parameters" in an equation.

A variable is a letter that stands for a number that can change. A parameter is a number, or a letter standing for a number, that does not change. "Constants" and "coefficients" in an equation are examples of parameters. In the equation: $Y = 100 + 2\,X^2$, Y and X are variables, 100 and 2 are parameters.

If I wanted to leave the exact values of the parameters undefined, I could write: $Y = a + bX^2$
Here a and b are parameters, not variables.

b) Functions expressing the expectations-augmented Phillips curve

"Expectations augmented Phillips curve" means relationships corresponding to either or both of these equations:

$$\pi_t = E\pi - \beta\,(u_t - u^*) + \epsilon_t$$
$$\pi_t = E\pi + \alpha\,(Y_t - Y^*) + \epsilon_t$$

In these equations, the parameters are α , β , u^* and Y^*. The variables are $\pi_t, E\pi, u_t$, Y_t , ϵ_t .

The subscript "t" indicates that the equations hold within a unit of time, a period. The period could be a month, a quarter (three months) or a year.

π_t is inflation in a period (month, quarter or average for a year.).

u_t is the unemployment rate in the same period.

Y_t is real GDP in the same period.

$E\pi$ is expected future inflation. It is the answer you would get if you went around to people in period t and asked: "what is your expected value for what inflation will be, not this year, but in future years after this year?" (Of course, you would first have to explain what "expected value" means, if the person you are asking has not take statistics.)

3

ϵ_t is a catch-all variable to account for everything else (anything other than real activity) that can affect inflation. This is the "supply shock" thing that I will get back to in the next section.

Though the two equations look a bit different, they are saying exactly the same thing. First I will explain the equation with u in it. Then I will explain the Y one.

c) The expectations-augmented Phillips curve with the unemployment rate u

In the equation $\pi_t = E\pi - \beta(u_t - u^*) + \epsilon_t$,

u^* is one particular value for the unemployment rate called the "natural rate of unemployment." It is also called the NAIRU, an acronym that stands for "Non-Accelerating-Inflation Rate of Unemployment." Other terms used for u^* are "full employment" or "maximum employment": the economy is at "full" or "maximum" employment when $u = u^*$.

Economists do not know what u^* is with certainty, but they can guess at it. Look at Figure 1 at the end of the section. It plots the actual unemployment rate in the four quarters of each year from 1949 through 2019, along with a commonly-accepted estimate (guess) of the natural rate of unemployment in the same years. See that in recessions (e.g. after 2008) the actual unemployment rate is higher than the natural rate of unemployment. See that u^* does change but it does so very gradually. It varies from decade to decade but it seems to have stayed in the range between 4 and 6 percent and *there is no long-term trend in u^**. It is about the same now as it was in the 1950s. (This is different from Y^*, which I will get back to in a minute.)

In an era when u^* is 6 percent, the equation would look like $\pi_t = E\pi - \beta(u_t - 6) + \epsilon_t$:

In an era when u* is 5 ½ percent, the equation would look like $\pi_t = E\pi - \beta(u_t - 5.5) + \epsilon_t$

The equation means that current inflation tends to equal expected inflation when the unemployment rate equals the natural rate. When the unemployment rate is bigger than the natural rate, inflation tends to be lower than expected inflation. When the unemployment rate is smaller than the natural rate, inflation tends to be higher than expected inflation. I say "tend to" because of the extra term ϵ, the supply-shock term. ϵ_t equals zero *on average* across many years, but in a particular period it can be positive (when special factors are pushing up on inflation) or negative (when special factors are reducing inflation).

Key point: the direction of causation in this equation is *from real activity to inflation, not the other way around*. That is, if you could do something that would increase u, that would tend to cause a decrease in inflation. The equations do *not* mean that you can increase u by decreasing inflation.

The parameter β gives the exact, quantitative relationship between $(u_t - u^*)$ and $(\pi_t - E\pi)$. β is often called the "slope" of the Phillips curve. To see why it is called the slope, subtract expected future inflation from both sides of the equation. That gives $(\pi_t - E\pi) = -\beta(u_t - u^*) + \epsilon_t$.
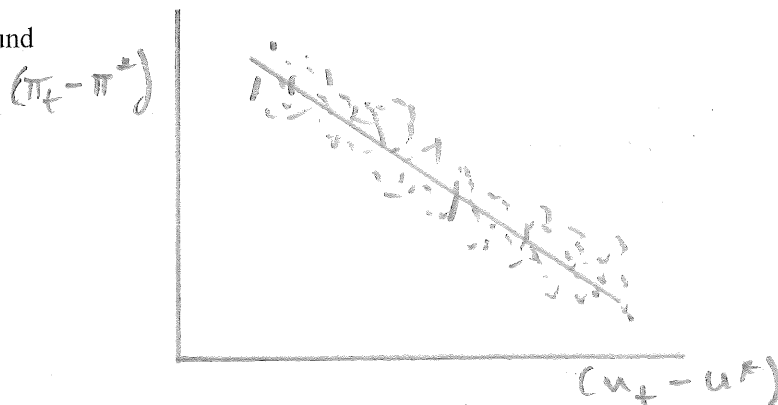
Now think about making a graph with $(u_t - u^*)$ on the horizontal axis and $(\pi_t - E\pi)$ on the vertical axis.

If you set ϵ equal to zero the graph would look like:

The slope is $-\beta$.
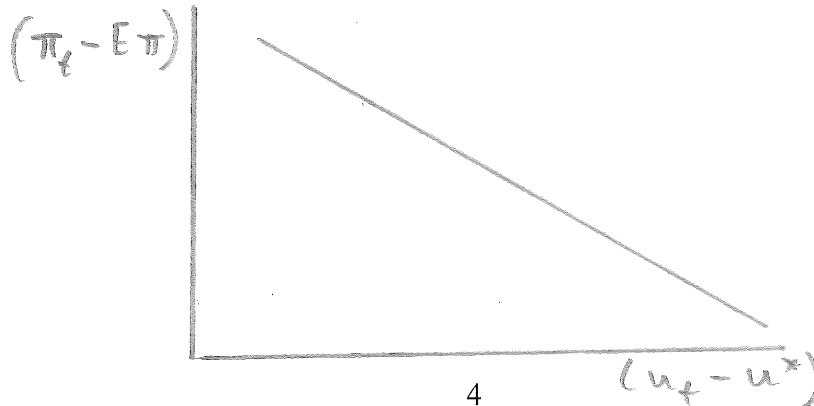
Because ϵ wriggles around
- it is sometimes zero,
  sometimes negative,
  sometimes positive
- the graph actually
looks like:

$(\pi_t - \pi^*)$

$(u_t - u^*)$

You might wonder why this relationship between inflation, expected inflation and the unemployment rate holds. I will get back to that later.

d) The expectations-augmented Phillips curve with real GDP Y

Now consider $\pi_t = E\pi + \alpha\,(Y_t - Y^*) + \epsilon_t$.

In this equation ϵ_t is the same "supply shock" as before. Y_t is real GDP this period. Y^* is a particular value for real GDP called the "natural rate of output." It is also called "potential output" or "potential real GDP." Y^* is the amount of real GDP the economy can produce when the unemployment rate is equal to u^*. The difference $(Y_t - Y^*)$ is called the "output gap.

The equation says that current inflation tends to equal expected future inflation when real GDP equals the natural rate. When real GDP is bigger than the natural rate, inflation is tends to be higher than expected inflation. When real GDP is smaller than the natural rate, inflation tends to be lower than expected inflation. α is called the "slope" of this expectations-augmented Phillips curve, because if you subtracted $E\pi$ from both sides, giving:

$(\pi_t - E\pi) = \alpha\,(Y_t - Y^*) + \epsilon_t$

and you made a
graph as before,
setting ϵ equal
to zero, you'd have:

$(\pi_t - E\pi)$

The slope is α.

$(Y_t - Y^*)$
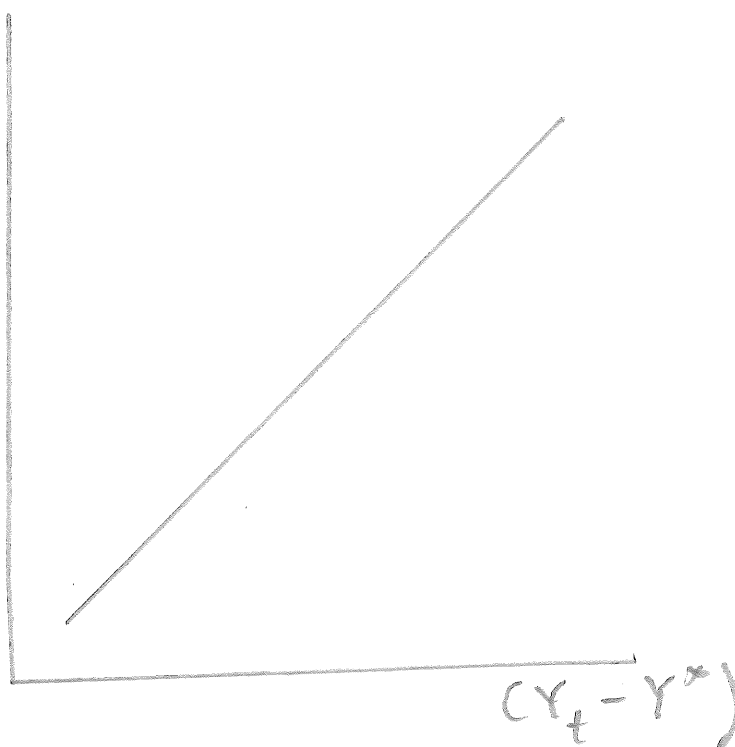
This equation and graph are entirely consistent with the expectations-augmented Phillips curve written in terms of u. They say just the same thing.

When $Y = Y^*$, $u = u^*$, by definition. Both equations say that realized inflation will (tend to) equal expected inflation.

When $Y > Y^*$, $u < u^*$. Both equations say that realized inflation will (tend to) be larger than expected inflation.

When $Y < Y^*$, $u > u^*$. Both equations say that realized inflation will (tend to) be less than expected inflation.

Of course, the coefficient α in the output version of the Phillips curve is not equal to β in the unemployment-rate Phillips curve. The magnitude of α relative to β depends on the exact quantitative relationship between fluctuations in real GDP and fluctuations in the unemployment rate. (From Econ 362, you might recall that this relatiosnhip is called "Okun's Law.")

e) Trend growth in Y^*

Like u^*, $Y^*$is something economists have to guess at. Figure 2 (at the end of the section) plots actual real GDP and an estimate of the natural rate. In recessions, actual real GDP is lower than the natural rate.

Notice that, unlike u^*, Y^* *grows every year.* This is because the labor force is always growing (at least in America), and technology is improving.

To understand this, think about two things you learning in Econ 362: the definition of the unemployment rate, and the aggregate production function.

The unemployment rate, expressed as a fraction rather than a percent, is:

u = (Number of unemployed people) / (Number unemployed + number employed)

The number on the bottom of the fraction, that is *(Number unemployed + number employed),* is called the "labor force." The labor force is the number of people who want to work, which includes those current working (employed) and those not currently working but looking for work (unemployed). The number of people in the labor force is determined by the size of the population, the fraction of the population that is of working age, and the fraction of the working-age populartion that wants to work - to "participate" in the labor force.

Given the unemployment rate, the number of people working depends on the size of the labor force. For an example, suppose that the unemployment rate is 10 percent. That means the number of employed people is 0.90 times the labor force. If the labor force is 100 million, that means 90 million people employed. At the same unemployment rate, a larger labor force means more people working. If the labor force is 110 million, 10 percent unemployment means 91 million employed.

When more people are working, the country produces more goods and services. As the labor force grows, the same natural rate of unemployment u^* means more people are working. Thus Y^* grows as the labor force grows.

But labor-force growth is not the *only* reason Y^* grows. The exact relationship between real GDP and the number of people working is given by the aggregate production function: $Y = AK^{\alpha}L^{(1-\alpha)}$. In this equation, Y is real GDP and L is the number of people working. When L gets bigger, Y gets bigger, of course. But two more variables in the aggregate production function are A and K.

A is a catch-all variable that accounts for the effect of improvements in technology. Better technology allows a country to produce more goods and services - real GDP - from any given inputs of resources. Thus, technological improvement as well as labor force growth causes Y^* to grow.

6

A also allows for changes in the average education of the labor force. Better-educated workers are more productive. Thus, an increase in the average education level might be another reason for growth in Y^*. One of the readings refers to this as "labor quality."

K is the aggregate capital stock in the country - machines, buildings, software, roads, bridges.... When investment spending is especially high, the capital stock grows relative to the labor force. Thus, another thing that affects Y^* is the amount of investment spending. Investmment spending is not always high enough to make capital grow faster than the labor force, but when it is, it is another thing that makes Y^* grow.

4) Controlling inflation through the Phillips curve

Imagine that you could directly control the level of real activity (*u* and *Y*) in the economy. And you have a target inflation rate, the way the Fed does. Let π^T denote your target inflation rate. For an example, suppose that your target inflation rate π_T is equal to 2 percent, like the Fed's. How do you keep inflation at your target?

First, guess what ϵ will be - that is, whether there are any upcoming special factors that will affect inflation in the upcoming period. For simplicity, assume that you guess ϵ will be zero (no special factors coming), which is true on average.

Next, find out what $E\pi$ is. Recall that $E\pi$ is expected future inflation, the public's expected value for what inflation will be in future years. The Fed finds out what $E\pi$ is mainly by looking at results of surveys that ask people - ordinary people, businesspeople, economic forecasters - what they expect future inflation will be.

Once you know what $E\pi$ is, compare it with your target π^T. The levels that you choose for real activity depend on whether $E\pi$ is greater than, less than or equal to π^T.

Suppose you have found $E\pi$ is equal to π^T. For a target equal to 2 percent, that means the expectations-augmented Phillips curve equations are:

$$\pi_t = 2 - \beta (u_t - u^*)$$
$$\pi_t = 2 + \alpha (Y_t - Y^*)$$

Looking at the equations, you can see that inflation π_t will equal 2 if $u_t = u^*$, $Y_t = Y^*$. That is what you want. So you set $u_t = u^*$, $Y_t = Y^*$.

What if $E\pi$ is greater than 2 percent, say 3 percent? That means the expectations-augmented Phillips curve equations are:

$$\pi_t = 3 - \beta (u_t - u^*)$$
$$\pi_t = 3 + \alpha (Y_t - Y^*)$$

In this case you *don't* want $u_t = u^*$, $Y_t = Y^*$: that would make π_t equal 3 - too high! What you want is $u_t > u^*$ and $Y_t < Y^*$ - you want the economy to be in a recession. Then realized inflation π_t will be less than 3. Of course you don't want to overdo it. But if you get it just right you can make $\pi_t = \pi^T$. (The exact values for u_t and Y_t that will make π_t equal 2 depend on the values of α and β.)

Finally, if $E\pi$ is less than 2 percent, say 1 percent, the expectations-augmented Phillips curve equations are:

$$\pi_t = 1 - \beta\,(u_t - u^*)$$
$$\pi_t = 1 + \alpha\,(Y_t - Y^*)$$

In this case you want real activity to be *above* the natural rate, that is $u_t < u^*$ and $Y_t > Y^*$ - a boom (the opposite of a recession), to make realized inflation π_t greater than 1. Again, the exact values for u_t and Y_t that will make π_t equal 2 depend on the values of α and β.

Now the Fed cannot control real activity like this. But it *can* affect real activity through interest rates. This is where the IS curve comes in.

D) The IS curve

1) Review and introduction

I hope you remember
the IS curve from
Econ 362. If you don't,
then education is wasted
on you. The IS
curve is:
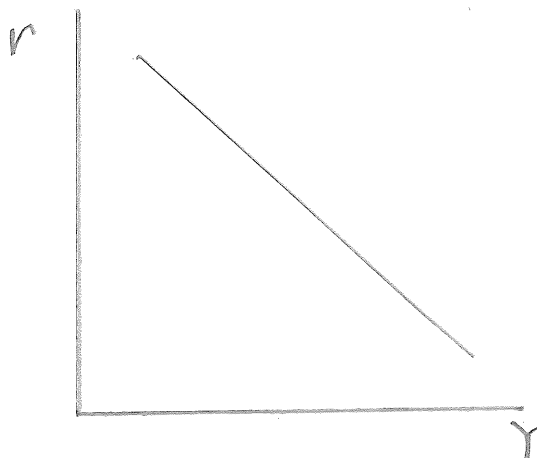
The horizontal axis is real GDP Y. The vertical axis is the real interest rate r. The real interest rate is the nominal interest rate i minus expected future inflation: $r_t = i_t - E\pi$.

The IS curve slopes down because, other things equal, an increase in the real interest rate reduces investment spending. (You may have learned about some additional reasons for the downward slope.)

In Econ 362, i was "the" nominal interest rate in an economy. In this class we know there is not just one nominal interest rate in an economy, there are many: long-term and short-term; loans *versus* yields on bonds; Treasury bonds *versus* corporate bonds and so on. The interest rates that matter for investment and consumption are interest rates on corporate bonds and loans to businesses and households. Thus the "i" that goes into the "r" for the IS curve must be a sort of average of interest rates on corporate bonds and loans to businesses and households. But, as I mentioned above, in this section we will assume for simplicity that the i_t that goes into the IS curve (as part of the real interest rate r_t) is the overnight rate.

2) The natural rate of interest

On the horizontal axis
of the IS curve, mark
the position of the natural
rate of output $Y*$. See that there is
one real interest rate which
makes real GDP Y equal $Y*$.
That real interest rate is called
the "natural rate of interest,"
denoted $r*$.
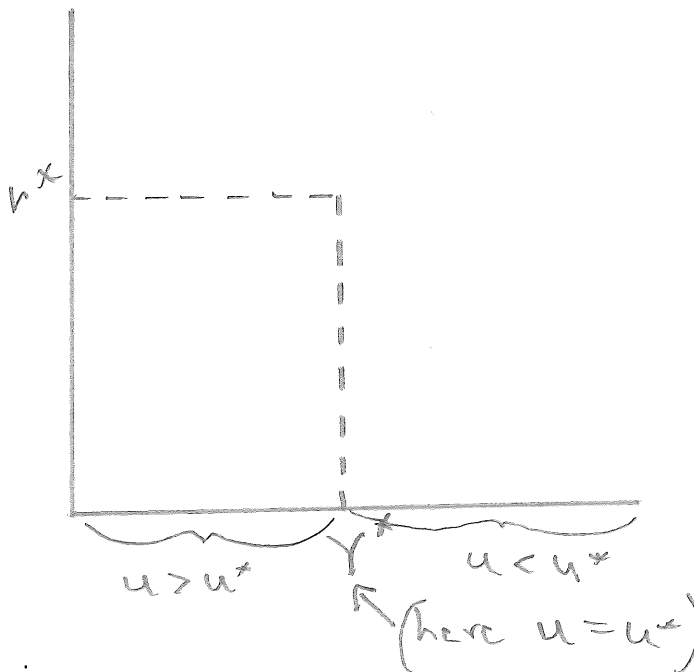
When $r_t = r*$, $Y_t = Y*$ and $u_t = u*$.

When $r_t > r*$, $Y_t < Y*$ and $u_t > u*$ - recession.

When $r_t < r*$, $Y_t > Y*$ and $u_t < u*$ - boom.

3) Controlling inflation through the real interest rate

Imagine that you could control the real interest rate r in the economy and you have a target inflation rate π_T. How do you keep inflation at your target? As before we will assume that ϵ in the expectations-augmented Phillips curve is zero (no special factors). Again, what you do depends on whether you believe $E\pi$ is greater than, equal to or less than your target π_T.

If $E\pi = \pi^T$ you want real activity to be at the natural rate, that is, you want $u_t = u*$ and $Y_t = Y*$. So you set $r_t = r^*$ .Then realized inflation π_t will equal your target.

If $E\pi > \pi^T$ you want $u_t > u*$ and $Y_t < Y*$, so you set $r_t > r^*$. Setting $r_t > r^*$ will cause a recession, making $u_t > u*$ and $Y_t < Y*$, which will push realized inflation down toward the target.

If $E\pi < \pi^T$ you want $u_t < u*$ and $Y_t > Y*$, so you set $r_t < r^*$. Setting $r_t < r^*$ will cause a boom, making $u_t < u*$ and $Y_t > Y*$, which will push realized inflation up toward the target.

E) Review: how to be an inflation-targeting central banker

Now let's put it all together. Imagine you are a member of a central bank policy committee that meets periodically (the FOMC meets about every six weeks) to choose a value for the interest-rate target i_t^T that will hold until the next time the committee meets. What does the committee do? First, you need to get information.

You also need to know the natural rate of interest $r*$. You can get estimates of $r*$ from central bank staff - central banks employ lots of economists to do things like that.

You compare $E\pi$ with your target π^T. That tells you whether you want the real interest rate r_t to be equal to $r*$ (to make $u_t = u*$ and $Y_t = Y*$), less than $r*$ (to make $u_t < u*$ and $Y_t > Y*$, a boom), or greater than $r*$ (to make $u_t > u*$ and $Y_t < Y*$, a recession).

But what about your *nominal* interest-rate target i_t^T?　　If you want $r_t = r*$, set $i_t^T = r*+E\pi$

If you want $r_t < r*$, set $i_t^T < r*+E\pi$

If you want $r_t > r*$, set $i_t^T > r*+E\pi$

F) More about expected future inflation $E\pi$

1) Introduction

As I have explained, the public's expected value for future inflation is an important number for central bankers. Where does it come from? How does the public think about future inflation? How do they form their expectaitons of future inflation?

2) "Anchored" versus "adaptive" expectations

Economics have observed that, over most of history, people seem to have formed their expectations of future inflation in one of two ways: "adaptive" expectations, or "anchored" expectations. In some historical eras, inflation expectations appear to have been adaptive. In other eras, inflation expectations appear to have been anchored.

Adaptive expectations means that people guess at what future inflartion will be by extrapolating from the recent past. A person says to herself "what was inflation last year?" (or perhaps on average over the last couple of years or so). And that number is what the person expects to continue in the future. $E\pi$ is affected by what inflation turned out to be last year.

Anchored expectations means that $E\pi$ is some stable number, *not* affected by what inflation turned out to be last year.

3) History of $E\pi$ in America

Figure 3 (at the end of the section) plots an estimate made by Federal Reserve economists of the public's expected value for future inflation in the PCE price index, starting in the first quarter of 1968 and ending in the second quarter of 2023. The estimate is based mainly on surveys of people who sell forecasts of inflation and other macroeconomic variables to firms. In each quarter, these "professional forecasters" are asked to guess what inflation will be in future years, on average. The plotted "Expected future" number for, e.g. 1976:4 (the fourth quarter of 1976) is the median answer the forecasters gave in that quarter. The figure also plots what PCE inflation had actually been over the year up to the quarter of the survey.

In the years from 1968 to the late 1990s, expectations were **adaptive**: see that $E\pi$ was unstable, and strongly correlated with what inflation had actually been over the past year.

Starting in the late 1990s, exectations became **anchored**. $E\pi$ was always close to 2 percent, no matter what actual inflation had been over the previous year. Actual realized inflation often turned out to be different from 2 percent, but this did not seem to affect the public's expected future inflation.

XIV

Federal Reserve people believe that expectations became anchored in the late 1990s because the public became aware that the Fed was trying to keep inflation equal to 2 percent, and came to believe that the Fed could hit that target *on average* - not all the time, not every year, but on average over several years. From the 1990s up to 2012, everyone knew the Fed had an inflation target of 2 percent but the Fed did not officially state it. In 2012, the FOMC made an ofcicial, explicit announcement that they had an inflation target and it was 2 percent in the PCE price index. The FOMC did that to make sure that expected inflation *stayed* anchored at 2 percent. Central bank policymakers *want* expectations to be anchored and equal to the inflation target.

4) What central bank policymakers want $E\pi$ to be

Central bank policymakers *want* $E\pi$ to be anchored and equal to the inflation target π^T.
They don't want $E\pi > \pi^T$ because that would require them to cause a recession (see above).
They don't want $E\pi < \pi^T$ because of the effective lower bound. I'll get back to this below.

G) Monetary policy when $E\pi = \pi^T$

1) Introduction

When $E\pi$ is anchored and close to 2 percent, the Fed's inflation target, the Fed wants to keep the real interest rate equal to the natural rate of interest. It sets the target (nominal) overnight rate i^T accordingly, so that equal to $i^T = r* + E\pi$. You might think that means the Fed would hold i^T stable. Not so! $r*$ **is a moving target.** Keeping $r = r*$ means making frequent adjustments to i^T to make r follow along with changes in $r*$. $r*$ can change because of changes in the natural rate of unemployment and output, and/or because of shifts in the IS curve.

2) Changes in $r*$ due to shifts in the IS curve

a) On a graph

Anything that shifts the IS curve up/out without (immediately) changing $Y*$ raises $r*$. Anything that shifts the IS curve down/in without immediately changing $Y*$ lowers $r*$.
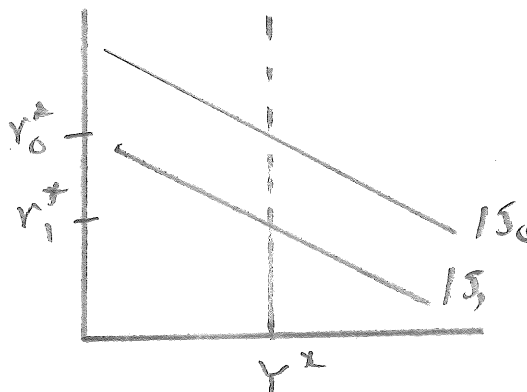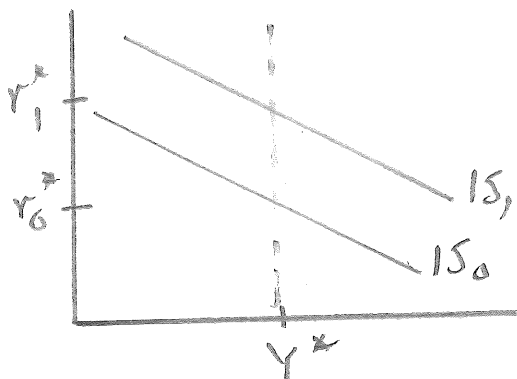
b) Fiscal policy

One thing that shifts the IS curve is "fiscal policy." In Econ 362 you learned that the IS curve is shifted out (in) by increases (decreases) in government purchases of goods and services, denoted G. Thus, increases (decreases) in G raise (lower) $r*$. You also learned that the IS curve is shifted in (out) by increases (decreases) in taxes-net-of-transfers T. Thus, increases (decreases) in T lower (raise) $r*$. Changes in G and/or T are called "fiscal policy."

c) Financial crises, stock market booms and crashes

11

A stock market crash (big fall in stock prices) reduces people's wealth and makes it harder for firms to raise funds by selling shares of stock. Thus, it reduces the amount of consumption and investment spending that would occur at any given real interest rate. That shifts the IS curve back/down, reducing r^*. A stock market boom (big increase in stock prices) has the opposite effect.

A financial crisis is a time when many financial intermediaries at once suffer liquidity crises. In a financial crisis many firms that could ordinarily issue bonds find that they cannot do so. Firms and households that cannot issue bonds find that it is harder to get loans from FIs. Thus, there is less investment and consumption spending, shifting the IS curve back/down, reducing r^*.

3) Changes in r^* due to changes in natural rates

a) Changes in u^*

Other things equal, increases in u^* tend to lower r^*. Decreases in u^* tend to lower r^*.

As an example, consider an increase in u^* from a low value u_0^* to a higher value u_1^*, say from 4 percent to 5 percent. An increase in u^* tends to decrease Y^* as it reduces the amount of labor input to the economy when $u=u^*$. Holding fixed the size of the labor force, the capital stock and technology, the economy produces less output when the unemployment rate is 5 percent than it can produce when the unemployment rate is 4 percent. Holding fixed the determinants of spending in the IS curve, this tends to raise r^*. Look at the two graphs below . The left-hand side graph shows an economy with the natural rate of unemployment equal to a low value u_0^* . The right-hand side graph shows the economy with a higher natural rate of unemployment u_1^*, and everything else the same. See that r_1^* is higher than r_0^*
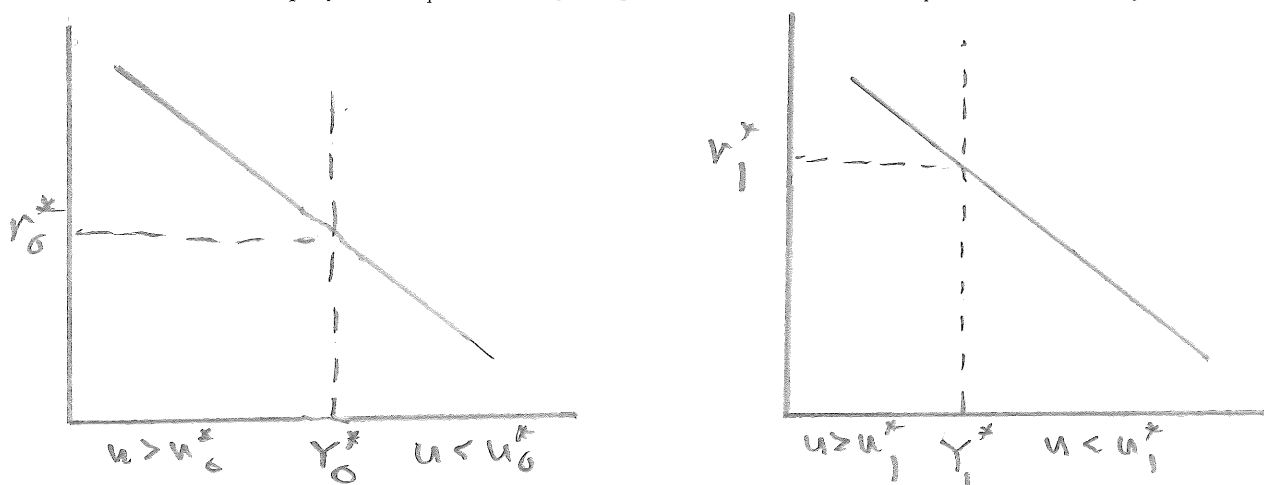
b) Changes in Y^* due to fluctuations in productivity growth

Earlier I said that Y^* grows every year, but it need not grow steadily. If productivity growth fluctuates, Y^* and r^* can change even if u^* remain fixed. For example, if productivity grows especially fast, that tends to increase Y^* and, holding the IS curve fixed, reduce r^*. The picture is simple so I won't draw it.

4) Accidental recessions and booms

a) Fed officials can be wrong about r^*

In my description of inflation targeting in subsections E) and G), I implicity assumed that Fed policymakers always know what the value of r^* is. If that is true, the only way that a recession can occur is for the Fed to cause one on purpose, because $E\pi > \pi^T$. Also, assuming that the Fed correctly maps out the IS curve and the expectations-augmented Phillips curve, it can always hit its inflation target if it wants to.

12

But it is not realistic to assume that Fed policymakers always know the true value of r^*. All that they have is an estimate of r^*, based on information prepared by Fed economists and their own judgement.

Let r^{*E} denote the FOMC's estimate of r^*. If the estimate is correct, that is if $r^{*E} = r^*$, then the situation is as I described it in subsections E) and G). But r^{*E} could be wrong! If $r^{*E} > r^*$, the FOMC has made an *overestimate* of r^*. If $r^{*E} < r^*$, the FOMC has made an *underestimate* of r^*. What happens in these cases?

b) Accidental recessions and booms and deviations of π from π^T

If r^{*E} is wrong, then there can be accidental recessions and booms - recessions and booms that the FOMC did not intend to happen. And that means inflation can accidentally deviate from the Fed's target.

Suppose $E\pi = \pi^T$, so that the Fed wants to make $u = u^*$, $Y = Y^*$. That means the FOMC wants $r = r^*$. So it will set $i^T = r^{*E} + E\pi$. The resulting real interest rate will be $r = r^{*E}$. If $r^{*E} = r^*$, the FOMC will get the result it wants - no recession or boom, inflation equal to the target.

But what if the FOMC has made an *overestimate* of r^*, so that $r^{*E} > r^*$? Then the value of i^T chosen by the FOMC will mean r is greater than r^*. There will be an accidental recession. Inflation will turn out to be too low, less than π^T.

On the other hand, if the FOMC has made an *underestimate* of r^*, so that $r^{*E} < r^*$, the value of i^T chosen by the FOMC will mean r is less than r^*. There will be an accidental boom. Inflation will turn out to be too high, higher than π^T.

Lots of information and lots of guesses go into an estimate of r^*, so there are lots of reasons the FOMC could be wrong about r^*. Here are a couple of examples. Try to make up some examples of your own!

c) Examples

### i) Fed economists are wrong about the position of the IS curve.

Suppose $E\pi = \pi^T$, so that the FOMC wants to make $u = u^*$, $Y = Y^*$. Fed economists guess that the IS curve is on the dotted line, but the IS curve is actually the solid line. Based on this wrong guess about the IS curve, the FOMC believes the natural rate of interest is r^{*E} (estimated r^*), but r^* is actually a lower number, as shown on the graph.
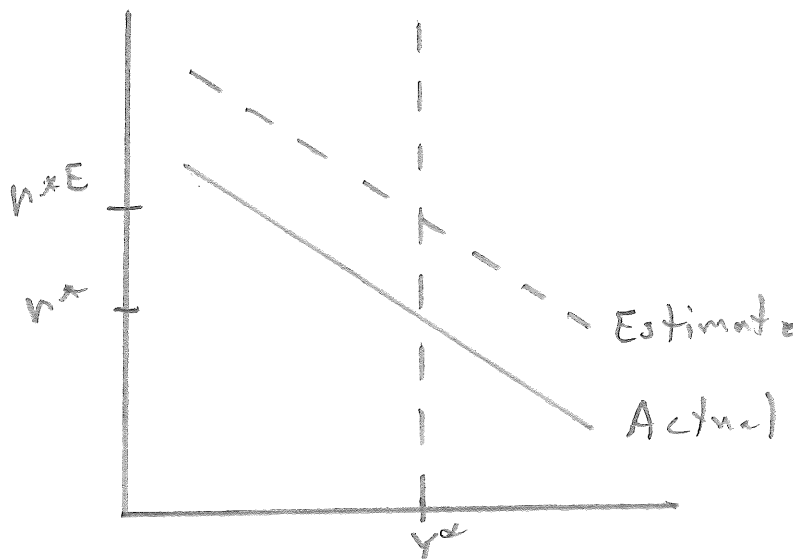
The FOMC will set the nominal interest rate equal to $i^T = r^{*E} + E\pi$. The resulting real interest rate will be higher than the true r^*. As a result, there will be an accidental recession - $u > u^*$, $Y < Y^*$ - and inflation will be too low - $\pi < \pi^T$.

ii) Fed economists are wrong about $u*$

Again $E\pi = \pi^T$, so that the Fed wants
to make $u = u^*$, $Y = Y^*$. Fed economists
guess that u^* is four percent, but $u*$ is
actually higher, equal to six percent.
As a result, Fed economists' estimate of
$Y*$ is too high. Based on these estimates,
the FOMC believes the natural rate of interest is
r^{*E} (estimated $r*$), but $r*$ is actually
a higher number, as shown on the graph.
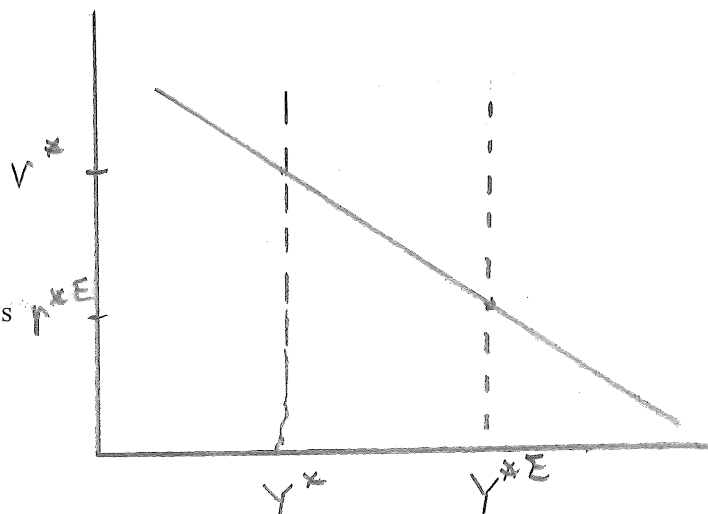
The Fed will set the nominal interest rate equal to $i^T = r^{*E} + E\pi$. The resulting real interest rate will be lower than the
true $r*$. As a result, there will be an accidental boom - $u < u^*$, $Y > Y^*$ - and inflation will be too high - $\pi > \pi^T$.

H) The lower bound problem

1) $i*$

The "lower bound" problem is a problem that can prevent Fed policymakers from stopping a recession, even if they know
the true value of $r*$. To talk about it, we will go back to our assumption that the FOMC knows the true value of $r*$. We
need one more piece of notation: $i*$. This is the the nominal interest rate that makes $r = r^*$. It is $i^* = r^* + E\pi$.
i* is sometimes called the "**neutral federal funds rate**" or the "**neutral rate**."
In terms of this new notation, we can restate the previous subsection as follows. When $E\pi = \pi^T$ the central bank will
want to set $i^T = i^*$; when $E\pi > \pi^T$ the central bank will want to set $i^T > i^*$; and so on.
Of course, in the previous subsection we assumed that once the central bank set a target, Fed staff could always push the
market rate i to hit that desired target. The problem is that.....

2) Sometimes $r*$ is negative

Sometimes the IS curve
shifts down/back so far that
$r*$ is a negative number.
This happened during the 2008
financial crisis. $r*$ appears
to have remained
negative for years after 2008.
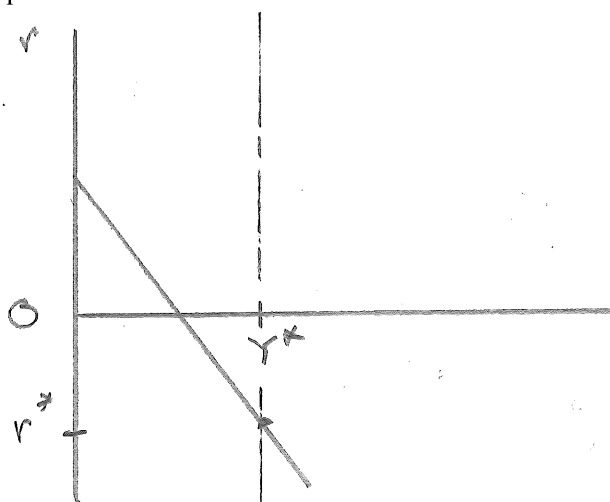
3) So i* can be negative

If $r*$ is a negative number, $i^* = r^* + E\pi$ might be a negative number. Of course, this depends on the value of $E\pi$ as well
as $r*$. Suppose that $r*$ is -3 percent and expected future inflation is 5 percent. Then the nominal interest rate that keeps r
$= r*$ is 2 percent. But what if expected future inflation is 2 percent? Then when $r*$ is -3 percent, the nominal interest rate
that keeps $r = r*$ is -1 percent! That can be a problem.

4) The lower bound on the nominal interest rate

a) Introduction

If $i*$ is negative, policymakers might not be able to make the market interest rate that low. The nominal interest rate can be pushed down only so far. The lowest possible value of the nominal interest rate is called the "lower bound" on the nominal interest rate.

b) The zero lower bound

Federal Reserve policymakers used to believe that the lower bound was zero. They called this the "zero lower bound" or "zero bound" on interest rates. They believed nominal interest rates on loans and yields on bonds, could not go below zero because there is a perfectly safe asset, easily available to anyone, which always pays a nominal interest rate exactly equal to zero. This asset is currency. Why would anyone make a loan at a negative interest rate, or buy a bond with a negative yield? (A negative yield means the market price of the bond is higher than the IOU.) That would be silly. You could just hold cash instead.

c) The effective lower bound

But this belief turned out to be wrong. In recent years several foreign central banks have in fact made market overnight rates negative. They did this by using the floor system and setting the interest rate paid on reserves at negative values - that is, they charge a bank interest on the funds the bank holds in its reserve account. The result has been that nominal interest rates are negative at the overnight maturity and even at longer maturities. (People believe that overnight rates will be negative for a while, so the expected value of the average value of overnight rates going out to some horizon is negative.)

It is still a bit puzzling why it works. Why would a bank hold funds in a reserve account at a negative interest rate - why not hold vaults full of cash instead? Why would anyone lend money at a negative interest rate - again, why not hold cash instead? One hypothesized explanation is that it is actually quite costly for a bank to hold vaults full of cash. Anyway, it *does* work. It is now clear that it *is* possible to make the market overnight rate go below zero. But Fed people believe it cannot be pushed *far* below zero. They believe it can only be pushed a little bit below zero - down to negative ½ percent, maybe. They call this lower bound, a little bit below zero, the "effective lower bound."

Unlike foreign central banks, the Fed policymakers have not set the target fed funds rate at a negative number, yet. But they have considered it.

d) The lower bound can prevent the Fed from keeping $r = r*$

The (effective) lower bound can prevent the Fed from keeping $r = r*$ when $r*$ is negative. As an example, suppose that $r*$ is -3 percent, expected inflation $E\pi$ is anchored at the Fed's target inflation rate π^T of two percent. Then $i^* = -3 + 2 = -1$. Suppose further than the "effective lower bound" is -1/2 percent. Then i^* is below the effective lower bound! To make $i = i^*$, the Fed would have to push i all the way down to -1. But it can't do that, because of the lower bound.

5) Result: recession

With the nominal overnight rate stuck at a too-high value, $i > r^* + E\pi$, so $r > r*$ so $Y < Y^*$, $u > u^*$. The economy is trapped in a recession. And because real activity is below the natural rate, inflation will tend to be lower than expected inflation. If expected inflation is anchored at π^T, realized inflation will be lower than the target. Everything is bad.

The U.S. was in this situation for many years after the 2008 financial crisis. During the crisis the FOMC cut the target overnight (fed funds) rate as low as they thought they could, "practically to zero." Figure 4 at the end of this section plots

the market overnight (fed funds) rate, monthly average from 1990 through March 2020. You can see that the overnight rate was practically zero from 2008 through 2015. But that was not low enough to make $r = r^*$. As you can see in Figures 1 and 3, through most of this time the economy was stuck in a recession - unemployment was higher than the natural rate - and realized inflation was below the 2 percent target,

6) A danger: expectations become unanchored

As realized inflation remained persistently below the 2 percent target, Fed policymakers feared that expected inflation would become "unanchored" and fall *below* 2 percent. It would have been *very bad* if expected inflation had become unanchored in this way. The lower bound problem would have become worse! Recall that the nominal interest rate that makes $r = r^*$ is equal to $r + E\pi$. If $E\pi$ started falling, i^* would have started falling further and further below zero. With the nominal interest rate stuck at zero, the real interest rate would have risen higher and higher above r^*, which would make Y fall lower and lower below Y^*, which would make inflation fall lower, which might make $E\pi$ fall lower and lower.....Economists call this a "deflationary spiral."

Fortunately, this did not happen after 2008. As you can see in Figure 3, expected inflation remained anchored to 2 percent.

7) An obvious fix: fiscal policy. But...

In Econ 362, you learned that fiscal policy can shift the IS curve. Anything that shifts the IS curve up/out increases r^*. So an obvious solution to the lower bound problem is fiscal policy: increase G or reduce T enough to push the IS curve up/out so that r^* is high enough so that $i = r^* + E\pi$ is above the lower bound.

Unfortunately, fiscal policy is not under the control of central bank policymakers. History suggests that the Federal government cannot be relied on to use fiscal policy to solve the lower bound problem. Look at Figure 5 at the end of this section, which plots G from 1947 to 2019. there was an upward trend in G over those years, as G kept up with growth in real GDP. What happened to G in the recession following the 2008 financial crisis? You can see that G was increased for a couple of years, but starting in 2010 G was actually cut even though unemployment was still very high, output was still far below the natural rate, and the natural rate of interest was still negative. (There were some decreases in tax rates in these years, but they were not sufficient to counteract the drag on spending from the decrease in G.) Thus, fiscal policy actually made the recession worse!

8) Solutions to the lower bound problem: changing the inflation target

a) Introduction

After 2008 economists came up with ways to avoid or at least mitigate the lower bound problem. To explain some of their ideas, I will have to introduce complications to the simple model we are using in this section. So I will save them for later. But two proposed solutions can be described in this section. They have to do with changes to the inflation target. So far, the Fed has not adopted either of them.

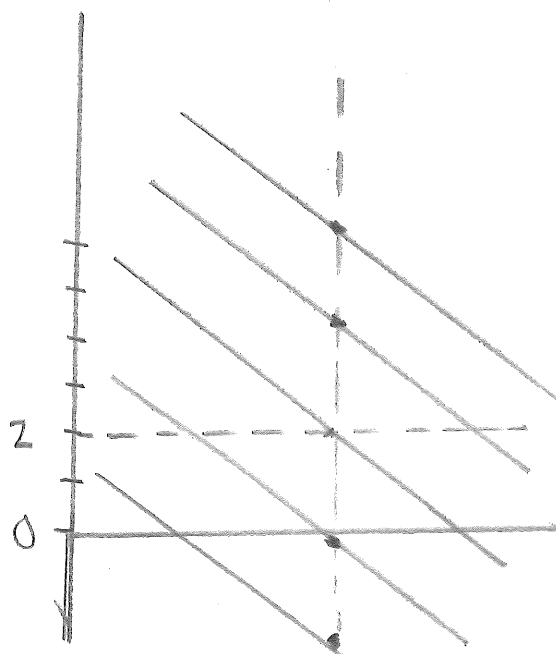### b) Raise the inflation target

One way to make it less likely that i^* will fall below the effective lower bound on the nominal interest rate would be to raise the inflation target π^T.

To see this, consider another example and another set of graphs. Start with the two IS curve graphs below. The left-hand side graph shows two economies, A and B. In *both* economies the long-run trend value of $r*$ is a low value, 2 percent, and the IS curves are wiggling around this trend due to short-term factors. The two graphs look exactly the same, of course:
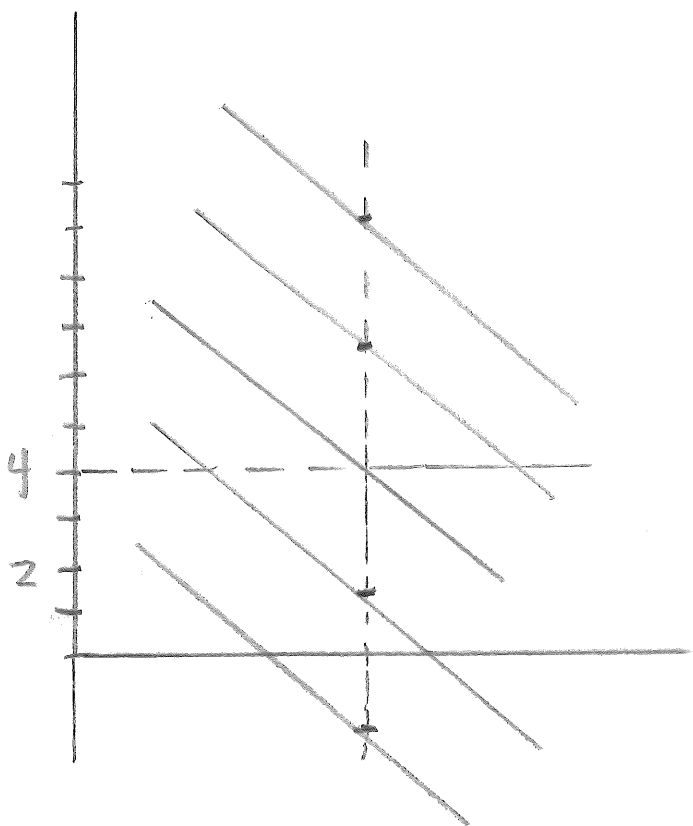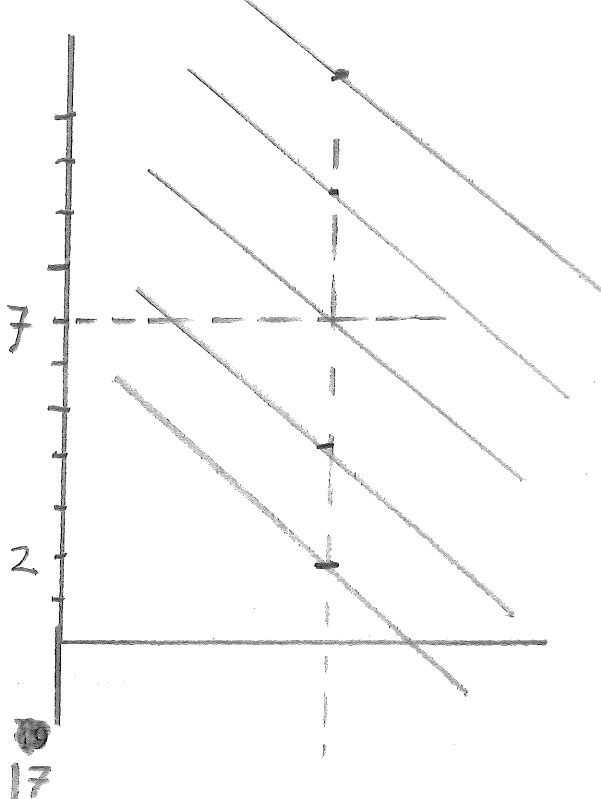
Economy A
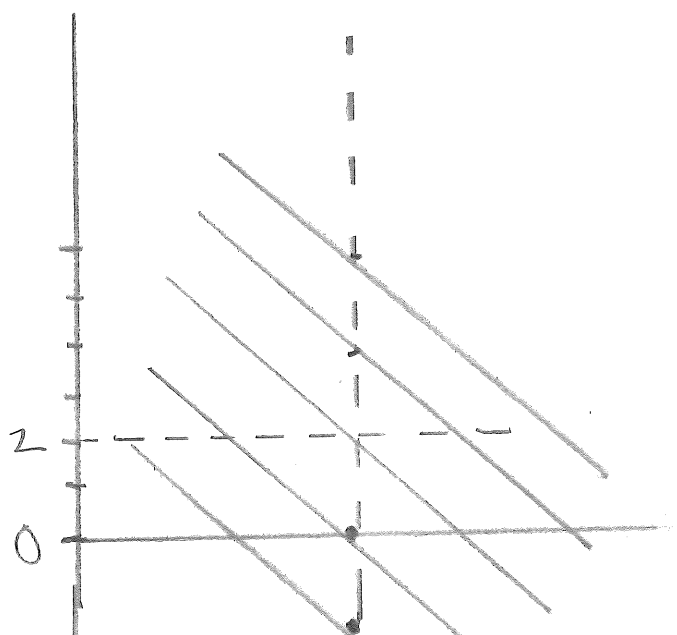
Economy B

But now consider the graphs that show $i*$. Suppose that in both economies expected future inflation is anchored and equal to the central bank's inflation target π^T. But the two economies have different inflation targets. In economy A the inflation target is 2 percent, as it is in the U.S. and most other countries. In economy B the inflation target is *higher*, equal to 5 percent. The two graphs below show what happens to $i*$ when the IS curve is shifting as shown above. See that $i*$ is *less* likely to be below zero in economy B, the one with the higher inflation target.

Economy A

Economy B

That seems easy enough! Based on this argument, many economists have recommended that the Fed (and other central banks) raise its inflation target above 2 percent, to help keep the effective lower bound from being a problem. So far, neither the Fed nor any other central bank has seriously considered doing this. The public seems to like the 2 percent number.

But I'm sure you can see that cutting the inflation target below 2 percent, say to zero, would make the lower bound problem *worse*. Thus, the lower bound is an important reason that Fed policymakers adopted an inflation target of 2 percent rather than zero.

c) Switch from inflation targeting to price level targeting

Economists have proposed that the central bank stop targeting inflartion and instead target a gradually-rising price *level*. Specifically, policymakers could say: "we will try to keep the value of the PCE price index on a prespecified path such that the price level target for any year is 2 percent higher than the target value for the prior year." This path would look like:
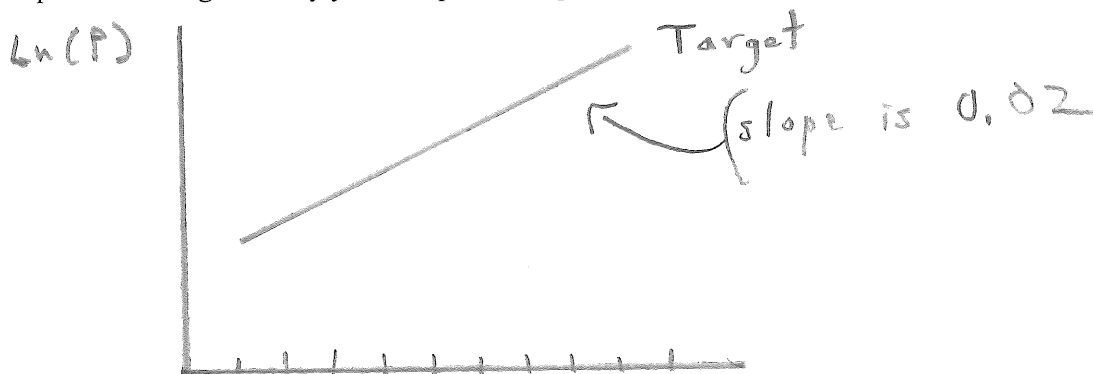

How is this different from targeting 2 percent inflation? Well, it *wouldn't* be different if the Fed could always hit its target. But if the Fed sometimes misses its target, targeting a rising price level is different from an inflation target in an important way. To see this, compare what happens when there is an accidental recession under inflation targeting *versus* price-level targeting.

Accidental recession under inflation targeting. Suppose the Fed is targeting 2 percent inflation. The public expects that the Fed will hit its target on average, so $E\pi = 2$. Expectations are anchored as they appear to have been after 2008, so that even when inflation turns about to be higher or lower than 2 percent in a given year, $E\pi$ remains 2 percent. Then there is an accidental recession because Fed staff overestimate r^*. Realized inflation falls below 2 percent. As soon as Fed economists realize what has happened, they need to set i^T equal to $i^* = E\pi + r^*$. Since $E\pi$ remains 2 percent, $i^* = 2 + r^*$. If r^* is less than -2, the lower bound may prevent them from setting i^T low enough, as you already know.

Accidental recession under price level targeting. Now suppose the Fed is targeting an upward-sloping path for the price level which grows 2 percent every year. The public's expected value for the price level in a future year is the Fed's price level target for that year. The public knows the actual price level may be higher or lower because the Fed may miss its target, but the expected value for the future price level is the Fed's target path.

Again suppose there is an accidental recession because Fed staff overestimate r^*. Realized inflation falls below the 2 percent increase in the price level set by the increasing price level target. So the actual price level follows the path of the dotted line:
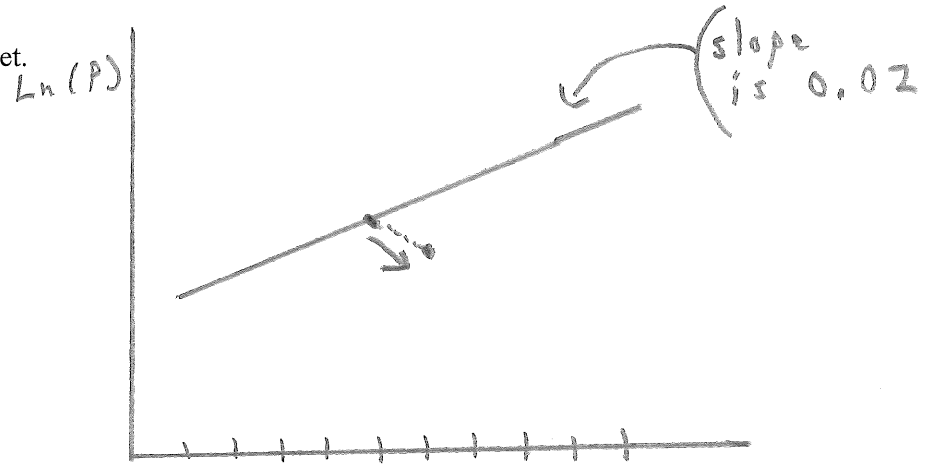
To get back to the planned path, what the price level has to do next year is this:
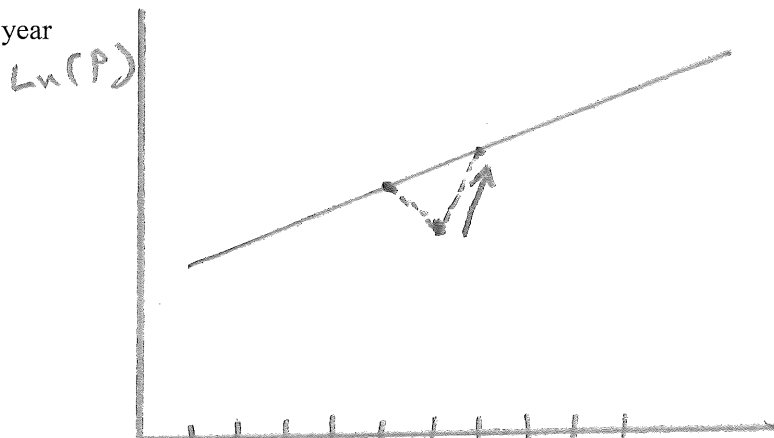
which is an increase of *more* than 2 percent. And the public's expected value for the price level is that it will get back to the Fed's target path. So expected future inflation is now *more* than two percent. With expected inflation greater than 2 percent, i^* is less likely to be below the lower bound.

The general point is that under price level targeting, a realized shortfall in inflation due to an accidental recession will automatically tend to increase $E\pi$, which raises i^* at any given r^*.

Figure 1 Unemployment rate and natural rate of unemployment (NAIRU)
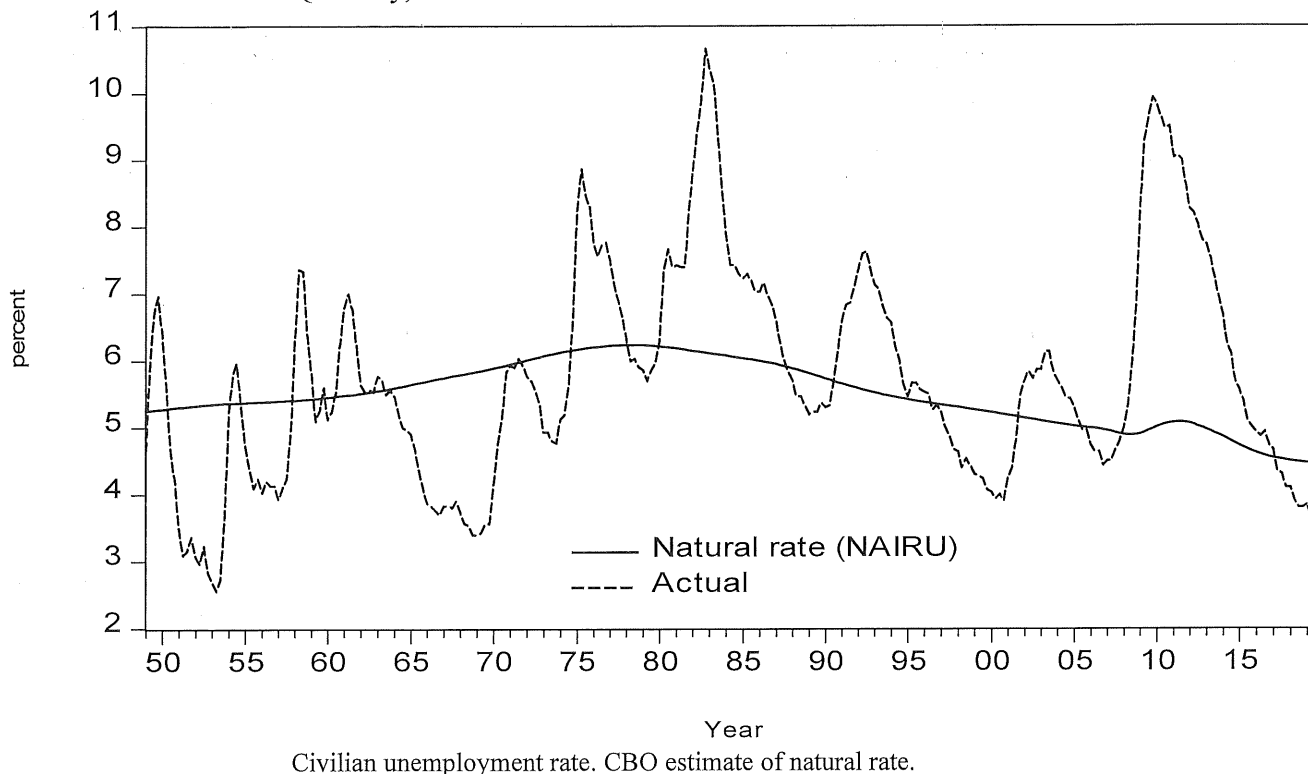Quarterly, 1949:1 - 2019:4

Year

Civilian unemployment rate. CBO estimate of natural rate.

Figure 2 Real GDP and natural rate of output (potential real GDP)
Quarterly, 1949:1 - 2019:4

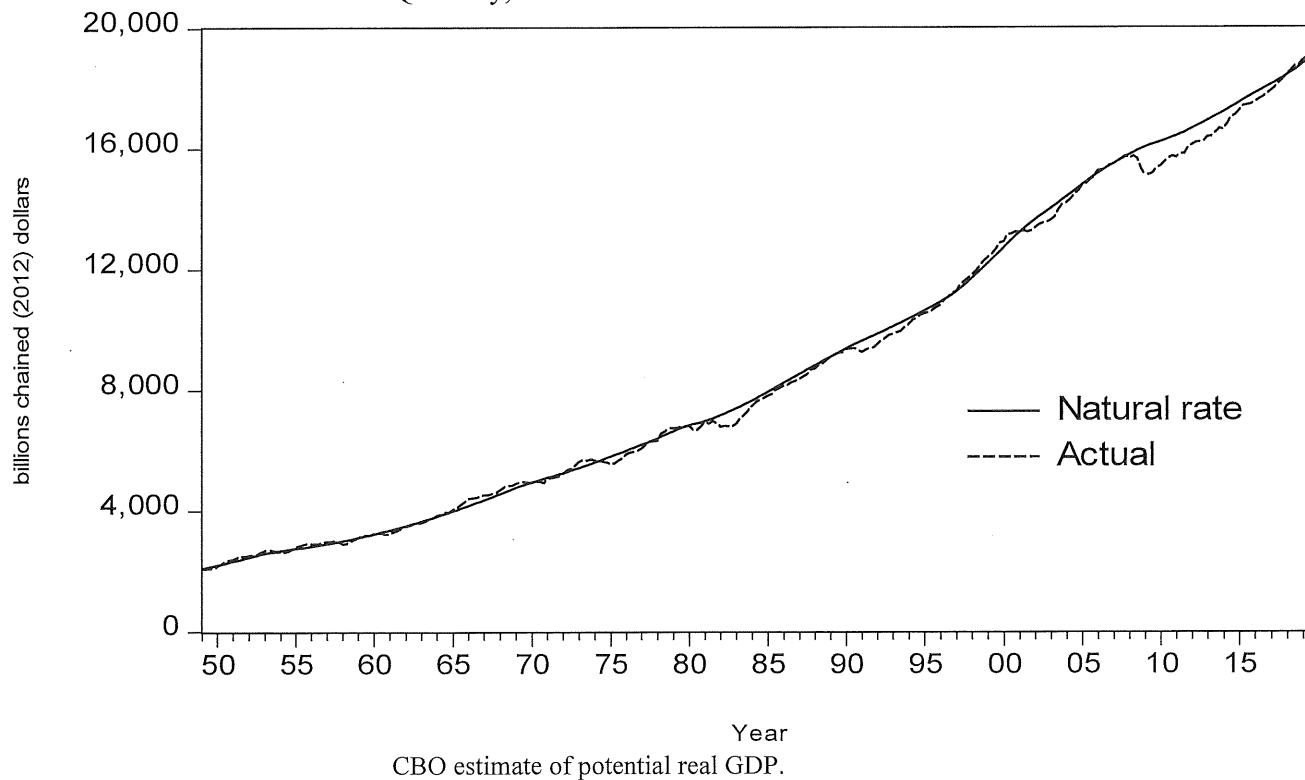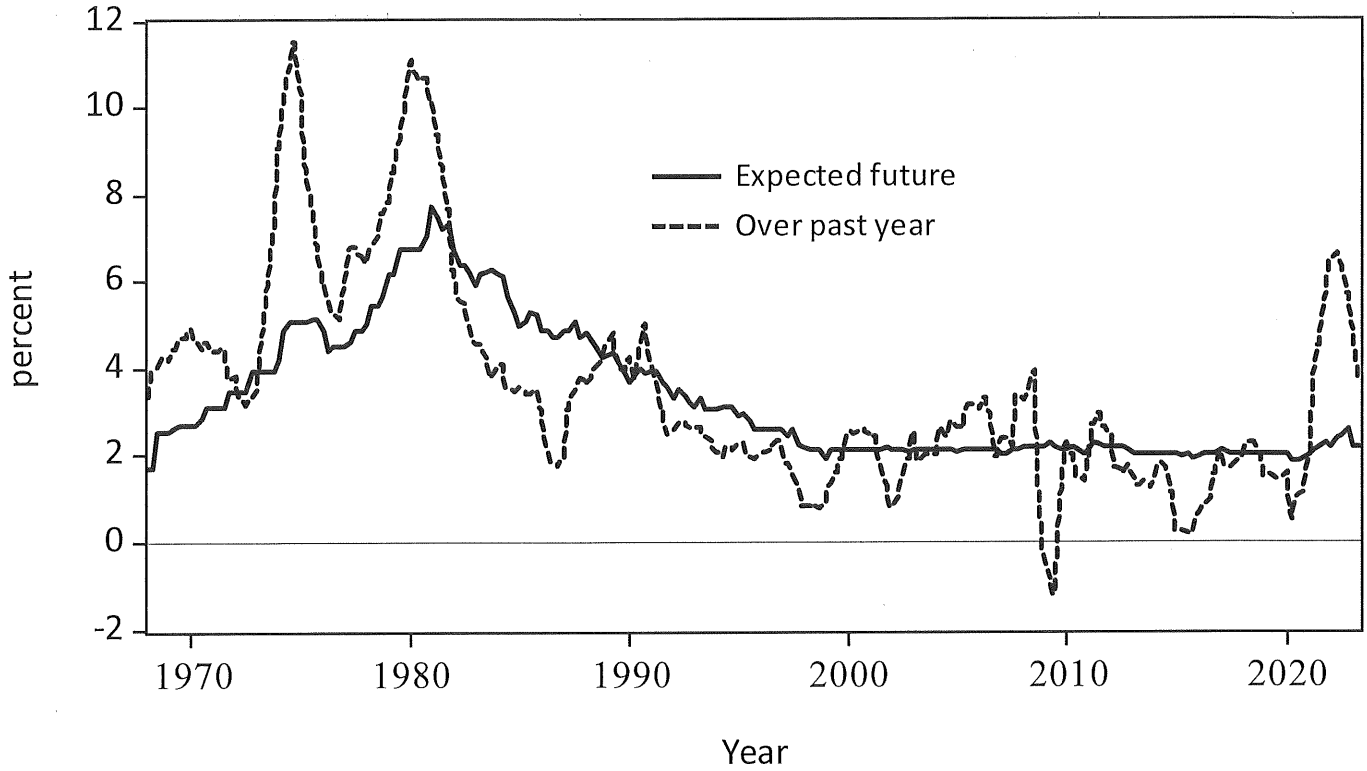Year

CBO estimate of potential real GDP.

20

Figure 3 Expected future inflation in PCE price index and actual current inflation
Quarterly, 1968:1 - 2023:2

Expected future is average over next ten years, Federal Reserve FRB-US model (series PTR) linked to Suervey of Professional Forecasters. Over past year is percent change from same quarter previous year.

Figure 4 Market federal funds rate (overnight rate)
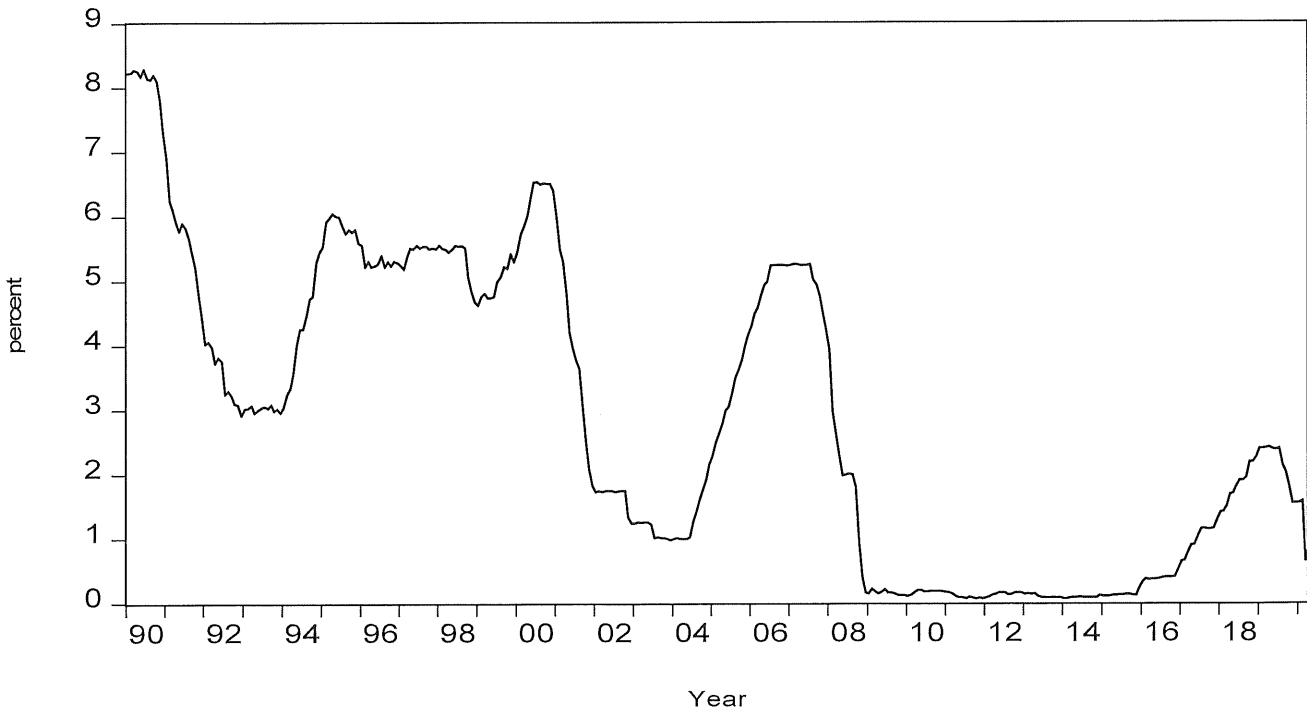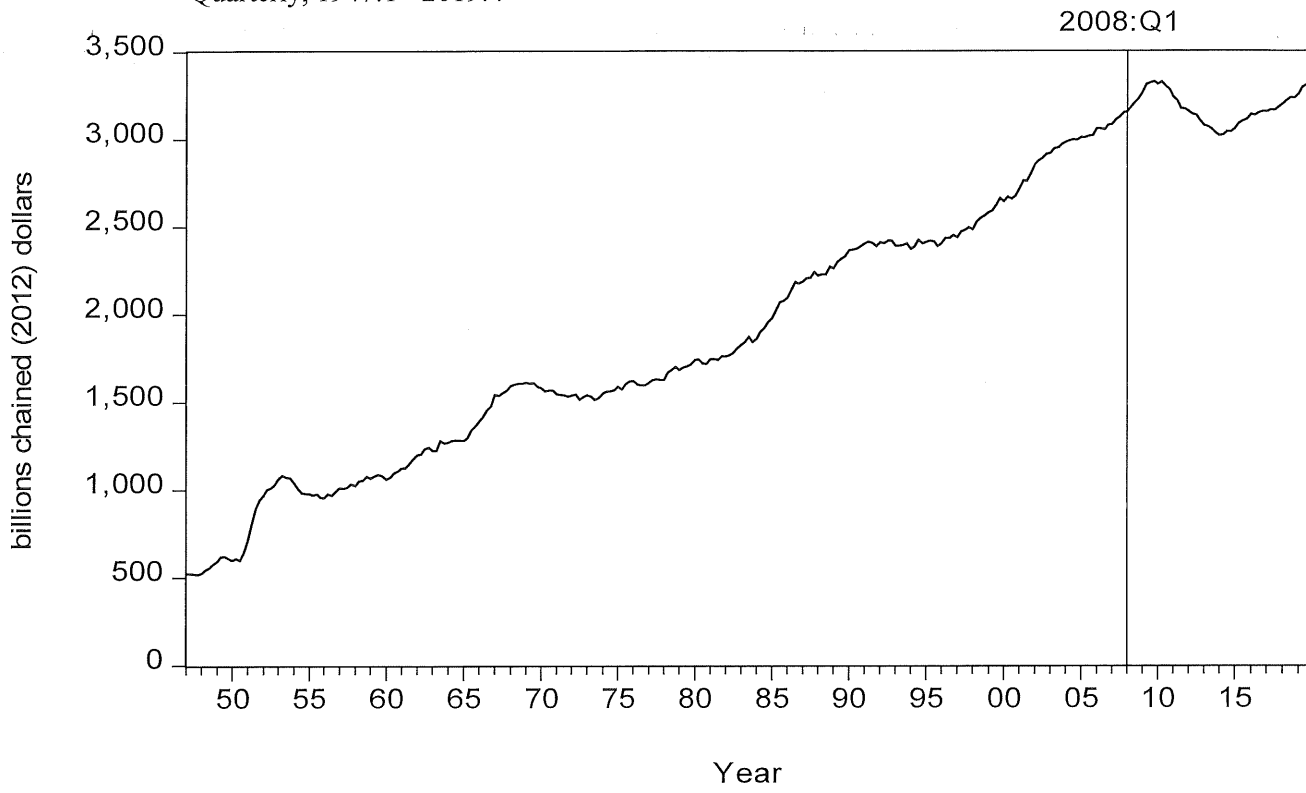Monthly average, January 1990 - March 2020

Figure 5 Real government purchases of goods and services
Quarterly, 1947:1 - 2019:4

©**Christopher Hanes 2020**   **Economics 450   Class notes**

XV) Monetary policy and macroeconomics, complications
    A) Introduction
    B) Long-term trend decrease in $r*$
        1) r* is lower now than it was before 2008
        2) High versus low trend value of $r*$
    C) Lags in the effect of monetary policy on real activity
        1) It takes time for a change in real interest rates to affect real activity
        2) Preemptive monetary policy
        3) More causes of accidental recessions and booms
    D) Supply shocks
        1) Introduction
        2) Where the expectations-augmented Phillips curve comes from
            a) Prices equal marginal cost plus a markup
            b) Determinants of marginal cost
                i) Wage inflation
                ii) Other determinants of marginal cost
            c) The resulting expectations-augmented Phillips curve
            d) Supply shocks go away after a while
        3) How do monetary policymakers respond to supply shocks?
            a) Introduction
            b) What they do if expectations are anchored
            c) What to do if expectations are adaptive
        4) Core inflation
    E) Real activity and inflation are affected by longer-term interest rates, not by the overnight rate
        1) Introduction
        2) What is the r that affects spending?
        3) "Unconventional monetary policies
            a) Introduction
            b) Term premiums and expected future target rates
            c) Forward guidance
            d) Quantitative reasing (QE)
        4) Long-term $E\pi$ and the lower bound problem
            a) Introduction
            b) If people are really smart, long-term $E\pi$ can be less than π^T
            c) $E\pi < \pi^T$ makes it more likely i will hit the lower bound
    F) The Fed's new (2020) policy strategy and Flexible Average Inflation Targeting (FAIT)

XV) Monetary policy and macroeconomics, complications

A) Introduction

In the previous section we made the unrealistic assumption that the real interest rate in the IS curve is the overnight real interest rate, so that the corresponding nominal interest rate is the one directly controlled by a central bank. In reality, the interest rates that affect spending, real activity and inflation are longer-term interest rates. In this section, we will deal with that complication. We will also deal with several other complications: changes in the long-term trend value of $r*$, lags in the effect of monetary policy on real activity and inflation, and supply shocks (cost-push shocks).

Taking account of those complications, we can discuss more problems policymakers faces and more techniques they use, such as "quantitative easing" and "forward guidance." We will also be able to discuss a recent modification the FOMC has made to its inflation-targeting regime, called Flexible Average Inflation Targeting.

B) Long-term trend decrease in $r*$

1) $r*$ is lower now than it was before 2008

In the previous section I talked about some things that can cause $r*$ to fluctuate from one year to the next. Economists believe there are also things that can change the *long-term trend* value of $r*$, that is the long-term level that $r*$ is fluctuating *around* from year to year. Furthermore many economists, including most Fed economists and policymakers, believe that in recent decades the long-run trend value of $r*$ has fallen: trend $r*$ is lower now than it was before the 2008 financial crisis, and it is likely to *remain* low for the foreseeable future.

Why has this happened? Economic theory suggests long-term trend $r*$ depends on the rate of growth in the labor force and the rate of improvement in productivity. Slower growth in the labor force and/or productivity tends to lower trend $r*$. Trend $r*$ can also be affected by changes in the supply of saving. A sustained increase in the supply of saving to an economy saving tends to lower trend $r*$. In the U.S., $r*$ may have fallen since 2008 (or starting even earlier) due to slower growth in the labor force (which definitely happened), slower productivity growth due to slower technological improvement (which may have happened), and an increase in the supply of savings from outside the U.S. (which probably happened).

Why do economists care about a decline in long-run trend $r*$? Because it makes the effective lower bound on i^T a bigger problem!

Suppose the long-term trend value of $r*$ is a big number. Then most downward temporary fluctuations in the IS curve will still leave $i*$ above zero. $i*$ will *rarely* hit the lower bound.

But what if the long-term trend value of $r*$ is a small number? Then temporary IS fluctuations will *often* push $i*$ below zero. $i*$ will *often* hit the lower bound.

2) High versus low trend value of $r*$

To see this in graphs, look below. The two upper graphs are IS curves, with r on the vertical axis, for two economies, A and B. In the left-hand side graph, for economy A, which the long-run trend value of $r*$ is 4 percent. The IS curve is wiggling around this due to short-term factors such as fiscal policy, stock market crashes and financial crises. In the right-hand graph, for economy B, the wiggles are the same size as in A but the trend value of $r*$ is lower, only 1 percent.
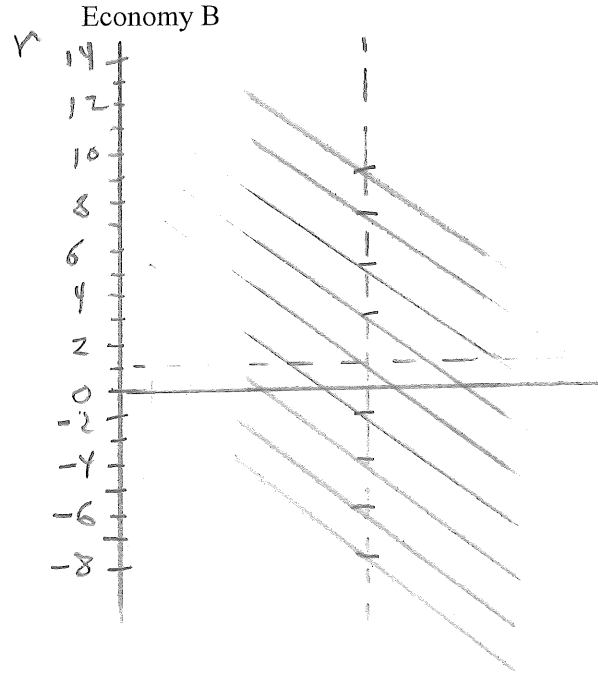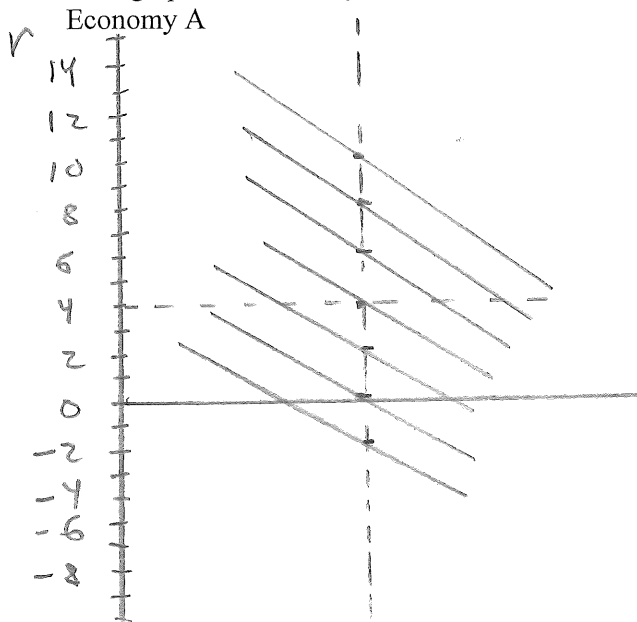
Economy A

Economy B

Now suppose that in both economies the target inflation rate π^T is 2 percent, and expected future inflation is anchored at this target: $E\pi = \pi^T = 2$. The two lower graphs show $i*$ in A and B. Remember that $i^* = E\pi + r*$. See that $i*$ is more likely to be below zero in economy B.
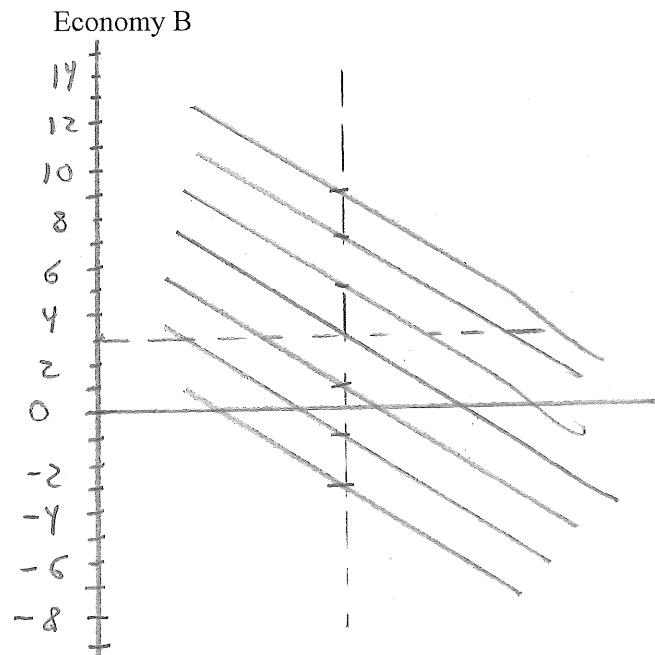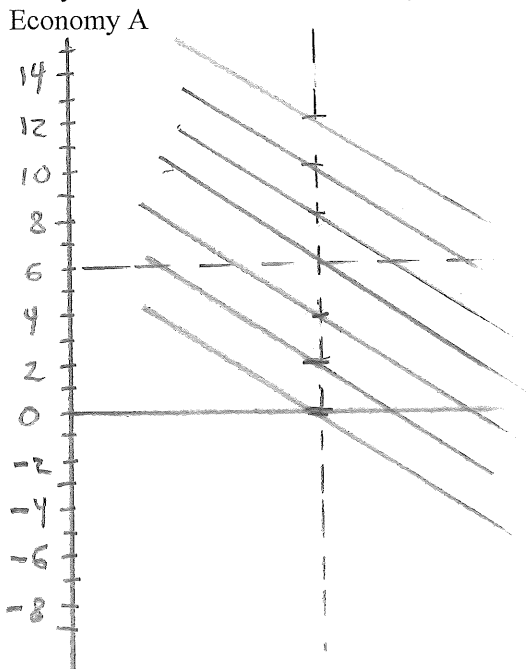
Economy A

Economy B

C) Lags in the effect of monetary policy on real activity

1) It takes time for a change in real interest rates to affect real activity

In section XIV) I assumed that changes in the real interest rate have an *immediate* effect on real activity. In fact, it appears to take time for changes in real interest rates to affect real activity. It seems to take six months to a year. Suppose, for example, that the Fed chooses randomly to hike its interest rate target this month, while expected future inflation remains stable, so that r goes up. Other things equal, Y will go down, u will go up. But not immediately; not this month or even next month. The effect on real activity will only begin to show up about six months from now, and will be at its maximum effect about a year from now.

2) Preemptive monetary policy

Because of the lag in effects of interest-rate changes, policymakers have to act "preemptively": that is, not just in response to *current* economic conditions but also in response to forecasts of *future* economic conditions. To do this, they need forecasts of what u^* and Y^* will be, and (in effect) where the IS curve will be, *six months to a year from now*, when today's value of the real interest rate will take effect.

In fact, before every FOMC meeting Federal Reserve economists create forecasts of what real activity and inflation will be a year in the future under various scenarios: if the FOMC makes no change to i^T between now and then; if the FOMC raises i^T at the upcoming meeting; if the FOMC cuts i^T at the upcoming meeting; and so on.

How would the FOMC use these forecasts? Here's an example. Suppose that everything in the economy is great right now: $\pi = \pi^T$, $E\pi = \pi^T$, and real activity is at its natural rates ($u = u^*$, $Y=Y^*$). All the FOMC wants to do is keep all this the same. But Federal Reserve staff forecast that "if interest rates remain at the current level, next year u will fall below u^* (which also means Y will rise above Y^*)." Acting preemptively, Fed policymakers would raise i^T *right now*, even though inflation is currently equal to π^T . Conversely, if the forecast was the opposite - "if interest rates remain at the current level, next year u will rise above u^* (which means Y will fall below Y^*)" - Fed policymakers would cut i^T right now even though inflation is currently equal to π^T .

A "symmetric" preemptive policy would mean that the FOMC responds to a forecast shortfall in real activity (a forecast that next year $Y<Y^*$, $u>u^*$) no more or less strongly than it responds to a forecast excess in real activity (a forecast that next year $Y>Y^*$, $u<u^*$). It tries to prevent future booms just as hard as it tries to prevent future recessions.

3) More causes of accidental recessions and booms

Because policymakers must set i^T in response to forecasts of *future* economic conditions, it is easy for accidental recessions and booms to occur, even if Fed policymakers have a perfect understanding of *current* economic conditions (which they don't, of course).

For example, suppose Fed policymakers set the current i^T based on a forecast that government purchases G will remain stable next year. But then Congress and the President unexpectedly push through a big increase in spending that pushes the IS curve out and raises r^*. That will cause an accidental boom and high inflation! The opposite happens if Congress and the President unexpectedly cut spending.

For another example, suppose Fed policymakers set the current i^T based on a forecast that government purchases G will *increase* next year (which will push the IS curve out and raises r^*). But then Congress and the President unexpectedly *hold spending stable*. That will cause an accidental recession!

3

The point is, there are lots of things that can cause the IS curve to shift or *not* shift unexpectedly between now and next year.

D) Supply shocks

1) Introduction

Go back to the expectations-augmented Phillips curve equations:

$$\pi_t = E\pi - \beta(u_t - u^*) + \epsilon_t$$

$$\pi_t = E\pi + \alpha(Y_t - Y^*) + \epsilon_t$$

I said the term ϵ_t in these equations, called the "supply shock" or "cost-push shock," is a catch-all variable to account for anything that can affect inflation *other* than real activity and expected future inflation. What is this about? And why does the expectations-augmented Phillips curve hold, anyway?

The exact mechanism that lies behind the expectations-augmented Phillips curve is a matter of debate among economists. But in this subsection I will tell you about a model of inflation that informs the thinking of many Fed economists. I will also tell you how central bank policymakers respond to supply shocks.

2) Where the expectations-augmented Phillips curve comes from

a) Prices equal marginal cost plus a mark-up

Think back to Econ 362. There (and in Econ 160, if you took it at Binghamton) you learned that a profit-maximizing firm selling in a perfectly competitive market will generally charge a price for a good (or service) equal to the marginal cost of producing one more unit of that good. A firm selling in a monopoly, oligopoly or monopolistically competitive market will charge a price equal to marginal cost *plus* a mark-up. The mark-up depends on factors including the elasticity of demand for one firm's output. Thus, the rate of inflation from one year to the next must be equal to some sort of average of changes in marginal cost *plus* changes in mark-ups across firms. Average mark-ups can perhaps vary somewhat from year to year, but not much. Thus, the rate of inflation is mainly determined by average changes in marginal cost.

b) Determinants of marginal cost

i) Wage inflation

Marginal cost is the cost of the "variable factors" needed to produce an additional unit of the good or service. Variable factors can include labor, fuel and materials. Labor represents a large share of marginal cost for most firms. Thus, an important determinant of most firms' marginal cost is the wages and salaries it must pay. And many materials and energy sources used by American firms are themselves produced in the U.S. The marginal cost of American-produced materials and energy is mostly the cost of American labor. That is American wages and salaries, again.

Averaging marginal cost across firms in an economy, one important determinant of price inflation must be the rate of wage inflation in the economy, that is, the average increase in wages and salaries from one period to the next. Denote wage inflation by π^W.

Many Fed economists believe that the effect of real activity on *price* inflation in the expectations-augmented Phillips curve largely reflects an effect of real activity on *wage* inflation. But why does real activity affect wage inflation?

The standard story is about the unemployment rate. When unemployment is high, an employer wants to cut the wages and salaries she pays below the level of wages that she expects other employers will be paying - what are her workers going to do, quit? Not when unemployment is high!

If all employers cuts wages at the same point in time, the average wage level would immediately plummet as employer A tries to cut the wages he pays below the wages paid by employer B while employer B is trying to cut the wages she pays below the wages paid by employer A....

However, all employers will *not* cut wages at the same point in time. Employers don't adjust their wage and salary schedules every week. Typically, an employer adjusts the wages and salaries of its workforce once a year. So as some employers are trying to cut the relative wages they pay, other employers aren't yet adjusting their wage schedules. This means that in response to an increase in the unemployment rate wages don't *plummet*, their rate of growth just slows down.

The opposite happens when unemployment is low.

Overall, you end up with a relationship between wage inflation and the unemployment rate that looks like:

$$\pi_t^W = E\pi^W - \beta(u_t - u^*)$$

ii) Other determinants of marginal cost

But labor is not the only determinant of firms' marginal cost.

One thing that matters is productivity growth. An increase in productivity tends to reduce the rate of growth in marginal cost relative to the rate of growth in wages.

Also, costs of materials and energy are affected by factors other than American wages. One is the price of oil, which is determined on the *world* oil market. Another is prices of food and other agricultural commodities, which are affected by bad harvests.

And of course, prices of imports are affected by changes in exchange rates. And so on, and so on.

Thus, one can describe price inflation this way: $\pi_t = \pi_t^W - Z + \epsilon_t$

In this equation, price inflation tends to be less than wage inflation due to trend productivity growth, represented by the parameter Z. ϵ_t represents effects of factors including fluctuations in the relative price of oil, food and other agricultural commodities, exchange rates, and so on.

c) The expectations-augmented Phillips curve

Putting it all together, price inflation is $\pi_t = E\pi - \beta(u_t - u^*) + \epsilon_t$
And that is the expectations-augmented Phillips curve with supply shocks!

d) Supply shocks go away after a while

An important thing about the supply-shock term ϵ_t is that it can be positive in some years (when special factors are pushing up on inflation) and negative in other years (when special factors are reducing inflation) but it averages to about zero over the long run. Oil prices, for example, rise much faster than wages in some years, and fall a lot in other years;

they do not keep rising fast (relative to wages) year after year. Harvests are good in some years, bad in others. Exchange rates go up and down. So when we have a big positive or negative supply shock in a year, you can be pretty sure that it is going to go away after a while.

3) How do monetary policymakers respond to supply shocks?

a) Introduction

Suppose that $E\pi = \pi^T$, so that the expectations-augmented Phillips curve becomes $\pi_t = \pi^T - \beta(u_t - u^*) + \epsilon_t$

Now suppose further that central bank policymakers know that a positive supply shock - a shock that tends to boost inflation - is coming or is underway. That means $\epsilon_t > 0$. If $\epsilon_t > 0$, the only way to keep inflation from rising above the target is for the policymakers to cause a recession, on purpose. Should they do that?

Conversely, if there is a negative supply shock coming, that is if $\epsilon_t < 0$, the only way for policymakers to keep inflation from falling below the target is to cause a boom. Should they do that?

Partly, it depends on whether the public's inflation expectations are *anchored* or *adaptive.*

If expectations are anchored, policymakers usually *don't* cause recessions and booms from affecting inflation.

If expectations are adaptive, they may think that they do need to cause recesssions and booms in response to supply shocks.

For an example, consider a positive supply shock $\epsilon_t > 0$.

b) What they do if expectations are anchored

In this case policymakers usually choose *not* to cause a recession. They try to keep $u = u^*, Y = Y^*$. The result is that in year t inflation rises above π^T. Inflation will remain high as long as the positive supply shock is hitting the economy. But remember that a supply shock goes away on its own after a while. So inflation will eventually return to target by itself, without the Fed having to cause a recession.

This is the strategy the Federal Reserve has followed at least since the late 1990s, when expectations became anchored.

c) What to do if expectations are adaptive

Adaptive expectations means that expected inflation is not stable. It changes in response to experienced inflation.

If expectations are adaptive and the economy is hit by a positive supply shock, the central bank pretty much *has* to cause a recession. If it doesn't, inflation will rise above target and remain too high *forever*. It won't go back down by itself.
To see this, consider a simple example of adaptive expections: $E\pi_t = \pi_{t-1}$. This equation says that in a period t the public's expected value for future inflation is equal to what inflation was in the previous period *t-1*. The expectations-augmented Phillips curve turns into:
$$\pi_t = \pi_{t-1} - \beta(u_t - u^*) + \epsilon_t$$
Let's say the supply shock goes away after just one year. What happens if the Fed tries to follow the same strategy I just described, trying to keep $u = u^*, Y = Y^*$, rather than causing a recession?

At first the result is similar to what happens under anchored expectations: in year t inflation rises above π^T, and inflation will remain high as long as the positive supply shock is hitting the economy.

But in this case inflation does *not* return to target by itself, even after the supply shock has gone away. That is because the experience of high inflation during the supply shock raises expected inflation $E\pi$ above π^T. In the year following the supply shock, call it *t+1*, the supply shock term $\epsilon_{t+1} = 0$. But because $E\pi_{t+1} > \pi^T$, $\pi_{t+1} > \pi^T$ as long as $u = u*, Y = Y*$. The only way for the central bank to *ever* get inflation back down again is to cause a recession!

Thus, the central bank is going to have to cause a recession *some* time - either in the year the supply shock hits or later - if it wants to get inflation back down to target. Might as well do it right away.

4) Core inflation

Generally, central bank policymakers need to know whether an observed change in inflation is it due to a supply shock, on the one hand, versus real activity and/or expected future inflation on the other.

One way to do this is to look at measures of wage inflation, which is not directly affected by supply shocks.

Another is to look at a measure of price inflation that strips out the effects of at least some supply shocks. The "core PCE" and "core CPI" are measures like that. They are indexes of consumer prices *excluding* energy and food. Thus, inflation in the core PCE price index or the core CPI are relatively unaffected by oil prices and harvests.

Note that stripping out food and energy prices does not strip out the effects of *all* supply shocks. During and after the covid crisis, there were a lot of weird supply shocks (see the readings).

E) Real activity and inflation are affected by longer-term interest rates, not by the overnight rate

1) Introduction

Back in section XIV), we assumed that the real interest rate in the IS curve is the overnight real interest rate. In reality, the interest rates that affect spending, hence enter into the IS curve, *longer-term* interest rates. That has two consequence we will discuss here. They both have to do with the lower bound.

First, when the overnight rate is at the lower bound, longer-term rates can still be above the lower bound, and a central bank can stimulate real activity if it can do something that reduces longer-term interest rates. In fact, Fed people believe they have two tools to do that. They are called "forward guidance" and "quantitative easing." That's good!

Second, because it is long-term, not short-term real interest rates that affect spending, the expected inflation term $E\pi$ that goes into the IS curve (as part of the relevant real interest rate) is correspondly long term. It is the public's expected value for the average inflation rate that will prevail going out five or ten years into the future. In a complicated way, that turns out to make the lower bound problem worse. That's bad.

2) What is the *r* that affects spending?

Think back to Econ 362. In that course, you probably talked about just one nominal interest rate, "the" nominal interest rate i. That meant there was just one real interest rate $r = i - E\pi$. The IS curve had this r on its vertical axis.

The IS curve was derived from a combination of two functions.

The "investment function" said that investment spending a negatively related to the real interest rate: $I(r)$.

The "consumption function" said that consumption spending was positively related to disposable income: $C(Y-T)$.

Putting the investment function and consumption function together with government purchases G gave the IS curve. The IS curve sloped down because of the negative effect of the real interest rate and investment spending. This effect was magnified by the "multiplier."

In this course we know that there are *many* nominal interest rates. Interest rates can be different at different maturities, for one thing. Let's focus on that now. For the nominal interest rate there is a corresponding real interest rate, which is equal to the nominal interest rate at that maturity minus the public's expected value for the average inflation that will prevail *over that maturity*. For example, if you are talking about a one-year rate, the corresponding real interest rate is the nominal one-year rate minus the public's expected value for inflation over the next year. If you are talking about a ten-year rate, the real interest rate is that rate minus the public's expected value for averageinflation over the next ten years. And so on.

So what is the maturity of the interest rates that go into the IS curve? The maturities that matter for the investment decisions of firms! Bonds and loans to finance new capital equipment, buildings, computers and such are about five to twenty years. Bills (such as commercial paper) sold to finance spending on inventories of goods to sell are of up to a year in maturity. And investment includes spending on construction of new houses and apartment buildings. Usually, such construction is ultimately financed by residential mortgage loans, which are typically twenty to thirty years.

Something that you might not have discussed in Econ 362, but which matters here, is that some forms of consumption spending are also financed by borrowing. Loans to finance new cars are about three years.

So, overall, what we've got is that there must be a lot of real interest rates, at different maturites, that affect spending and hence go into the IS curve. (In mathematical terms, you might say that the interest rate in the IS curve is really a *vector* of many interest rates.)

We don't want to get *too* complicated here. We will just say that "the" nominal interest rate that goes into the IS curve is an average or general level of medium- to long maturity interest rates. And the $E\pi$ that goes into the r for the IS curve is the public's expected value for the average inflation rate from now to about ten years from now.

3) "Unconventional" monetary policies

a) Introduction

When the overnight rate is stuck at the lower bound, a central bank can still stimulate real activity if it can reduce longer-term interest rates. The Fed's methods to do this are called forward guidance and quantitative easing. The Fed first began using these tools in the recession that followed the 2008 financial crisis. At the time they were called "unconventional monetary policies," as distinct from the conventional monetary policy of adjusting the target fed funds rate. Nowadays these tools have become become pretty conventional. They were used again during the covid crisis.

b) Term premiums and expected future target rates

From earlier in the course, recall that any interest rate longer than overnight can be described as the sum of three components: the expected value of the average value of the central bank's target overnight rate from the present out to the time the bond or loan matures; a term premium (present in all long-term bonds and loans); and any relevant default-risk and/or liquidity premiums (present in some bonds): $_n i_t = E\left[\dfrac{i_{Td} + i_{Td+1} + \ldots\ldots}{number\ of\ days}\right] + {}_n\tau + other\ premiums$

Let's pull the current target overnight rate out front, which gives us:

$$_n i_t = \frac{1}{number\, of\, days}\ i_{Td} + E\left[\frac{i_{Td+1} + \ldots\ldots}{number\ of\ days}\right] + {}_n\tau + other\ \ premiums$$

Suppose the current overnight-rate target i_{Td} is already on the lower bound. then the FOMC can't reduce longer-term interest rates by cutting the current value of i^T .

But it *can* try to push down on the term premium ${}_n\tau$. The Fed's method to do that is called quantitative easing.

It can *also* try to push down on $E\left[\dfrac{i_{Td+1} + \ldots\ldots}{number\ of\ days}\right]$. The Fed's efforts to do that are called forward guidance.

c) Forward guidance

"Forward guidance" means that the central bank tries to convince the public that it will keep the target overnight rate equal to zero for a long, long time.

An example of forward guidance would be an announcement by the FOMC that it will hold the target overnight rate to zero until at least a year from now. Another example of forward guidance would be an FOMC announcement that it will not raise the overnight rate above zero until the unemployment rate has gone down to five percent. Both of these types of forward guidance have been used by the Fed.

d) Quantitative easing (QE)

"Quantitative easing" is what most people call it. Fed people called it LSAPs. Yet another name for it is "balance sheet policies."

Whatever you call it, it means buying bonds in open-market operations, not to increase reserve supply but to reduce the supply of bonds avaiable to the public, and thus push down on term premiums. There is no generally-accepted theory to explain why this might work. But most Fed people believe it does work. In 2014 Ben Bernanke, then chair of the FOMC, said that " "The problem with QE is it works in practice but it doesn't work in theory."

4) Long-term $E\pi$ and the lower bound problem

a) Introduction

In section XIV) H) 8), I showed that an increase in the inflation target can alleviate the lower-bound problem, as it increases expected future inflation and hence increases $i*$. Of course, the same thing would hold in reverse: a decrease in the target would make the lower-bound problem worse as it reduces expected future inflation and reduces $i*$.

As usual, I was assuming that expected future inflation was equal to the target inflation rate. What if there was no change in π^T, but expected future inflation fell below π^T ? That would also lower $i*$ and make the lower bound problem worse. In section XIV) H) 6) I discussed such a situation.This was a situation where the economy had been stuck in a recession for a while, with realized inflation persistently falling below π^T , and inflation expectations became unanchored. Essentially, because people see that the Fed has been persistently missing it's inflation target, expectations switch from being anchored to being adaptive.

The problem I am talking about now is kind of like that, but more complicated. It is specifically about long-term $E\pi$. It is something that can be true all the time, not just after the economy has been stuck in a recession for a long time.

b) If people are really smart, long-term $E\pi$ can be less than π^T

Sometimes i^* is above the lower bound. Suppose that when it is, Fed policymakers seek to keep π equal to π^T, and are able to hit the target on average. This is what we have said so far.

But sometimes i^* is below the lower bound. Suppose that in this situation Fed policymakers will use unconventional monetary policies, but this probably won't be enough to keep $Y=Y^*$, so it is likely that there will be a persistent recession with $\pi < \pi^T$ and the Fed won't be able to stop it. This will be true even if inflation expectations remain anchored. (After all, this is exactly what happened in most years between 2008 and 2015, as shown in Figure 3 in section XIV.)

The result will be that across many years, with i^* sometimes below the lower bound, *average* inflation will be less than π^T.

Now suppose that the public is really smart, smart enough to figure all this out. Then long-term $E\pi$ will be less than π^T, *all* the time, even when the economy is not in a recession at all! Fed people often put it this way: given the lower bound problem, there may be a "downward bias in [long-term] inflation expectations" relative to π^T.

c) $E\pi < \pi^T$ makes it more likely i will hit the lower bound

We have just said that long-term $E\pi$ is the $E\pi$ that goes into the IS curve. So this phenomenon makes it more likely that i^* will fall below the lower bound on i at times when $r^* < 0$.

Well, this is a mess. It makes the effective lower bound an even bigger problem!

F) The Fed's new (2020) policy strategy and Flexible Average Inflation Targeting (FAIT)

Federal Reserve policymakers have been very worried about the lower bound problem, especially because they think r^* is now low and likely to remain low. They have been even more worried because they think it is possible that long-term $E\pi$ can fall below π^T for the reason I just explained, bexause of the "downward bias" in long-term inflation epxectations, relative to π^T.

To alleviate this problem, in August 2020 the FOMC announced to the public that was changing its its inflation-targeting strategy. It was *not* raising the 2 percent inflation target, or adopting price-level targeting. But it *was* making other changes. One of these changes is called "Flexible average inflation targeting" (FAIT).

How is FAIT different from the inflation targeting strategy the FOMC announced back in 2012? What other changes were made in the FOMC's strategy?

That is for you to figure out, from the readings!